

108.0001

SECTOR \$
PATENT #3

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: Jeffrey L. Eakin

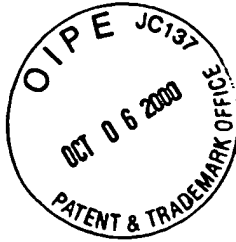
Serial No.: 09/619,411

Filed: July 19, 2000

For: METHODS AND APPARATUS FOR
PROCESSING AND DISTRIBUTING
INFORMATION RELATING TO
COSTS AND SALES OF PRODUCTS

Group: 2753

Examiner: Not Assigned Yet.



I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner of Patents and Trademarks, Washington, D.C. 20231, on the date set forth below:

Signed: Tlynthia Jordan

Name: Tlynthia Jordan

Date: October 4, 2000

Chapel Hill, North Carolina
October 4, 2000

Assistant Commissioner of Patents
and Trademarks
Washington, DC 20231

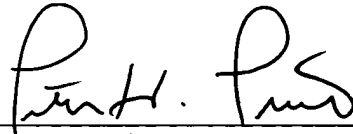
**RESPONSE TO NOTICE TO FILE MISSING PARTS OF NONPROVISIONAL
APPLICATION**

Sir:

Enclosed is a copy of the Notice to File Missing Parts of Nonprovisional Application and Declaration and Power of Attorney for the above identified case.

Also enclosed is our check in the amount of \$130 to cover the fee.

Respectfully submitted,

A handwritten signature in black ink, appearing to read "Peter H. Priest", written over a horizontal line.

Peter H. Priest
Reg. No. 30,210
Law Offices of Peter H. Priest
529 Dogwood Drive
Chapel Hill, NC 27516
(919) 942-1434

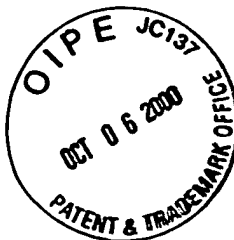


UNITED STATES PATENT AND TRADEMARK OFFICE

COMMISSIONER FOR PATENTS
UNITED STATES PATENT AND TRADEMARK OFFICE
WASHINGTON, D.C. 20231
www.uspto.gov

APPLICATION NUMBER	FILING/RECEIPT DATE	FIRST NAMED APPLICANT	ATTORNEY DOCKET NUMBER
09/619,411	07/19/2000	Jeffrey Lynn Eakin	108.0001

Peter H Priest
Law Offices of Peter H Priest
529 Dogwood Drive
Chapel Hill, NC 27516



FORMALITIES LETTER



OC000000005394773

Date Mailed: 09/13/2000

NOTICE TO FILE MISSING PARTS OF NONPROVISIONAL APPLICATION

FILED UNDER 37 CFR 1.53(b)

Filing Date Granted

An application number and filing date have been accorded to this application. The item(s) indicated below, however, are missing. Applicant is given TWO MONTHS from the date of this Notice within which to file all required items and pay any fees required below to avoid abandonment. Extensions of time may be obtained by filing a petition accompanied by the extension fee under the provisions of 37 CFR 1.136(a).

- The oath or declaration is missing.
A properly signed oath or declaration in compliance with 37 CFR 1.63, identifying the application by the above Application Number and Filing Date, is required.
- To avoid abandonment, a late filing fee or oath or declaration surcharge as set forth in 37 CFR 1.16(e) of \$130 for a non-small entity, must be submitted with the missing items identified in this letter.
- The balance due by applicant is \$ 130.

*A copy of this notice **MUST** be returned with the reply.*

Customer Service Center

Initial Patent Examination Division (703) 308-1202

PART 2 - COPY TO BE RETURNED WITH RESPONSE

10/11/2000 WABRHAH1 00000006 09619411

01 FC:105

130.00 OP

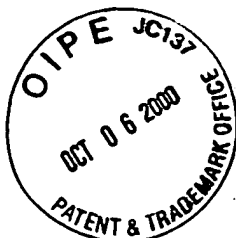


UNITED STATES PATENT AND TRADEMARK OFFICE

COMMISSIONER FOR PATENTS
UNITED STATES PATENT AND TRADEMARK OFFICE
WASHINGTON, D.C. 20231
www.uspto.gov

APPLICATION NUMBER	FILING DATE	GRP ART UNIT	FIL FEE REC'D	ATTY. DOCKET NO	DRAWINGS	TOT CLAIMS	IND CLAIMS
09/619,411	07/19/2000	2761	708	108.0001	25	21	3

Peter H Priest
Law Offices of Peter H Priest
529 Dogwood Drive
Chapel Hill, NC 27516



FILING RECEIPT



OC000000005394772

Date Mailed: 09/13/2000

Receipt is acknowledged of this nonprovisional Patent Application. It will be considered in its order and you will be notified as to the results of the examination. Be sure to provide the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION when inquiring about this application. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. If an error is noted on this Filing Receipt, please write to the Office of Initial Patent Examination's Customer Service Center. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the PTO processes the reply to the Notice, the PTO will generate another Filing Receipt incorporating the requested corrections (if appropriate).

Applicant(s)

Jeffrey Lynn Eakin, Residence Not Provided;
Christopher Huyette Howe, Residence Not Provided;
Joseph Eric Lipovich, Residence Not Provided;
Mark Wesley, Residence Not Provided;
Michael James Platt, Residence Not Provided;

mac KENZIE

Please
Correct.

Continuing Data as Claimed by Applicant

Foreign Applications

If Required, Foreign Filing License Granted 09/12/2000

Title

Methods and apparatus for processing and distributing information relating to cost and sales of products

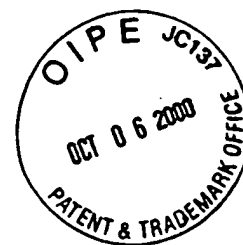
↓
COSTS

Preliminary Class

705

IN THE UNITED STATES
PATENT AND TRADEMARK OFFICE

Declaration and Power of Attorney



As the below named inventors, we hereby declare that:

Our residence, post office address and citizenship are as stated below next to our names.

We believe we are the original, first and joint inventors of the subject matter which is claimed and for which a patent is sought on the invention entitled **METHODS AND APPARATUS FOR PROCESSING AND DISTRIBUTING INFORMATION RELATING TO COSTS AND SALES OF PRODUCTS**, the specification of which was filed on July 19, 2000 as U.S. Serial No. 09/619,411.

We hereby state that we have reviewed and understand the contents of the above identified specification, including the claims, as amended by an amendment, if any, specifically referred to in this oath or declaration.

We acknowledge the duty to disclose all information known to us which is material to patentability as defined in Title 37, Code of Federal Regulations, 1.56.

We hereby claim foreign priority benefits under Title 35, United States Code, 119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

None

We hereby claim the benefit under Title 35, United States Code, 119(e) of any United States provisional application(s) listed below:

None

We hereby claim the benefit under Title 35, United States Code, 120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, 112, we acknowledge the duty to disclose all information known to us to be material to patentability as defined in Title 37, Code of Federal Regulations, 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application:

None

We hereby declare that all statements made herein of our own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any

patent issued thereon.

We hereby appoint the following attorney with full power of substitution and revocation, to prosecute said application, to make alterations and amendments therein, to receive the patent, and to transact all business in the Patent and Trademark Office connected therewith:

Peter H. Priest

(Reg. No. 30,210)

Please address all correspondence to Peter H. Priest, Law Offices of Peter H. Priest, 529 Dogwood Drive, Chapel Hill, North Carolina 27516. Telephone calls should be made to Peter H. Priest by dialing Area Code 919-942-1434.

Full name of 1st joint inventor: Jeffry Lynn Eakin

Inventor's signature  Date: 9/25/00

Residence: Ridgefield, CT 06877
Citizenship: US
Post Office Address: 125 Eleven Levels Road
Ridgefield, CT 06877

Full name of 2nd joint inventor: Christopher Huyette Howe

Inventor's signature  Date: 9/25/00

Residence: Stamford, CT 06902
Citizenship: US
Post Office Address: Apt. S-302
1450 Washington Blvd.
Stamford, CT 06902

Full name of 3rd joint inventor: Joseph Eric Lipovich

Inventor's signature  Date: 9/25/00

Residence: Trumbull, CT 06611
Citizenship: US
Post Office Address: 11 Moosehill Rd.
Trumbull, CT 06611

Full name of 4th joint inventor: Mark Wesley Mac Kenzie

Inventor's signature Mark Wesley MacKenzie Date: 9/25/00

Residence: Brewster, NY 10509
Citizenship: US
Post Office Address: 81 Acorn Road
Brewster, NY 10509

Full name of 5th joint inventor: Michael James Platt

Inventor's signature Michael James Platt Date: 9/25/00

Residence: Fairfield, CT 06432
Citizenship: US
Post Office Address: 279 Lynnbrook Rd.
Fairfield, CT 06432

Db 2548 GGCCTGGGCGGGGCTCAGGCTTATATCCAGCATTTTGGAGGCGGGGTGA 2607
QY 805 TCACCTTAAGGTCAGGAGTTTGAAGCCAGCCTGCCCAACATGTGTAACGTTGCTCTACT 864
Db 2608 TCACCTTAAGGTCAGGAGTTTGAAGCCAGCCTGCCCAACATGTGTAACGTTGCTCTACT 2667
QY 865 AAAAATACAAAATTAGACAGGCGGTGGGCAACATCTGTAAATTCAGCTTACTCAGAG 924
Db 2668 AAAAATACAAAATTAGACAGGAGTGGGCGCACCTGTAAATCTAGCTACTCGGAG 2727
QY 925 GCTAACACAGAAATTCCTTGAACCTGGAGGAGGAGTTGACAGGCAATTCGACTC 984
Db 2728 GCTAGAGTAGAGAACTGCTTGCACCCAGAGGAGGTTGACGTGTGCACTGCACTC 2787
QY 985 CAGCCTGGGCAACACAGGAGACTCTTGTCTCAAAAAAAAAAAAA 1030
Db 2788 CAGCCTGGGTCACAAAGGAGACACTGTCAACAAATTAACAACA 2833

RESULT 15
US-10-027-632-137060/c
; Sequence 137060, Application US/10027632
; Publication NO. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 137060
; LENGTH: 541
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-137060

Query Match 19.2%; Score 199.4; DB 13; Length 541;
Best Local Similarity 81.9%; Pred. No. 1.6e-33;
Matches 245; Conservative 1; Mismatches 42; Indels 11; Gaps 1;

QY 745 GGCCTGGGCGGATGCTGATGCTGTGCTTGTCCAGTCTTTGGGGCCGAGGCAAGTGA 804
Db 407 GGGTGGGTGAGGTGCTCAGCCTGTAAATCCAGACCTTTGGAGGCTGAGGCAAGTGA 348
QY 805 TCACCTTAAGGTCAGGAGTTTGAAGCCAGCCTGCCCAACATGTGTAACGTTGCTCTACT 864
Db 347 TCATTTGAGACCAAGAGGTTGAGACCAAGCTTGGCCAACTGTGTAACCCGCTCTACT 288
QY 865 AAAAATACAAAATTAGACAGGCGGTGGGCAACATCTGTAAATTCAGCTTACTCAGAG 924
Db 287 AAAAATACAAAATTAGACAGGAGTGGGCAACATCTGTGTAACCCGCTCTACTCAGAG 228
QY 925 GCTAACACAGAAATTCCTTGAACCTGGAGGAGGAGTTGACAGTGAAG----- 973
Db 227 GCTGAAGCAGAGAAATCACTTGAACCCGAGAGGTGAGGTTGACGTGAGCAAGATTGCA 168

QY 974 CCATTGCACTCCAGGCTGGGCAACACAGTGAAGACTTGTCTCAAAAAAAAAAAAA 1032
Db 167 CCACTGCACTCCAGGCTGGGCAACAGAGTGAAGACTTGTCTCAAAAAAAAAAAAA 109

Search completed: March 6, 2005, 08:03:45
Job time : 635 secs

Db	248	TCACCTGAGGTCAGGAGCTTTGAACACAGCTGGCCAAATATGCTAAACCCCACTTCTACT	189					
Qy	865	AAAAATATCAAAATTTGACACAGGCGTGTGTGCACACATCTGTAAATTCAGTACTACAGAG	924					
Db	188	AAAAAATACAAAATAAGACCGAGGCGTGTGTGCAGGACACTGTATCCAGTACTTGGAG	129					
Qy	925	GCTAACACAGGAATAATTCCTTGAACCTGGGAGGCAAGGTGGACGTGAG-----	973					
Db	128	GCTGAGGCGAGGAGATTCGCTTATCTCTGGAGGCAAGGTGTGAGTAACCGAGATCACA	69					
Qy	974	CCATTGCACCTCCAGCGCTGTGGCAACACAGTGAGACTTGTGTCAAAAAAAAAAAAAA	103					
Db	68	CCATTGCACCTCCAGCGCTGTGGCAACAGATGAAGCTTGTGTCAAAAAAAAAAAAAA	9					
RESULT 2								
LOCUS	BM907568/c	1024 bp	mRNA linear EST 12-MAR-2002					
DEFINITION	AGENCOUNT 6641246 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:5735872							
ACCESSION	BM907568	5', mRNA sequence.						
VERSION	BM907568.1	GI:19357947						
KEYWORDS	EST.							
SOURCE	Homo sapiens	(human)						
ORGANISM	Homo sapiens							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.							
AUTHORS	1 (bases 1 to 1024)							
TITLE	NIH-MGC http://mgs.nci.nih.gov/.							
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)							
COMMENT	Unpublished (1999)							
Contract: Robert Strusberg, Ph.D.								
Email: cga@sbs-remail.nih.gov								
Tissue Procurement: DCTD/DIR/gazdar								
cDNA Library Preparation: Life Technologies, Inc.								
DNA Sequencing by: Agencourt Bioscience Corporation								
Clone distribution: MGC clone distribution information can be								
found through the I.M.A.G.E. Consortium/LNL at:								
http://image.lnl.gov								
Plate: LHAM12743 row: h column: 17								
High quality sequence stop: 518.								
Location/Qualifiers								
1. 1024								
/organism="Homo sapiens"								
/mol_type="mRNA"								
/db_xref="taxon:9606"								
/clone="IMAGE:5735872"								
/tissue_type="large cell carcinoma"								
/lab_host="DH10B (phage-resistant)"								
/clone_lib="NIH MGC 68"								
/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.8 kb. Library constructed by Life Technologies."								

Query Match	18-9%	Score 196.6	DB 5	Length 1024
Best Local Similarity	66.2%	Pred. No. 9.2e-22		
Matches 304	Conservative	0	Mismatches 144	Indels 11
			Gaps	1
QY	586 AATTTTATTATATATAAGATGTGATGAGTACTACAAACACGAAAAATATA	645		
DB	461 AAGATTAATATATTATATAATAGACTACGTGTGATGTTACCAACGATAGAAAGCA	402		
QY	646 CAGATTTTGTGTCGCTTTTGCTAAATGTTCCTGACAAAATCTTAGCCAGTCTTCA	705		
DB	401 CAGGTAAAGATAGCTTTCTCTCTCCATACCAAGTAAATTATACCTAACAGATTGGGC	342		
QY	706 TTTTGGTTTGAGATGAAGACTACTATTAGTTTAAAGCAGGGGCTGGGCGATAGCTGANG	765		
DB	341 AATCTTAGCTATGAAAAATATACTTAAAGATTTTCTTAGCCGGGACATGTGGCTCAC	282		

QY 766 CCTGTGGTCCAGAGGCTTTGGCCGGGGCCAGAGGAGGTGATCACTTAAGTCCAGAGTTTG 825

Db 281 CCTTAAATCCACAGACTTTTGGAGGCCAGAGGGGGCGGATCACTCGAAGTCCAGAGTTTG 222

QY 826 AGACCAAGCCTTGCCCAACATGATGTAAACGTTGTCTCTACTTAAAAATACAAAAATTAGACAG 885

Db 221 AGACCAAGCCTTGCAACATGATGTAAACCTCGATTTCTACTTAAAAATACAAAAATTAGCAG 162

QY 886 GCGTGGTGGCAGACATCTGTAATTCACGTTCTCAGTACTCAGAGAGGCTAACACAGAAAAATTCCTT 945

Db 161 GTGTGGTGGCATGTGCTCTTAATCCACGCTACTCAGAGAGGTGAGACAGAGAAATTCCTT 102

QY 946 GAACTTGGAGGACAGAGTTTGCAGTGA-----GCCATTGCACTCCAGCCTGGGC 994

Db 101 GAACTTGGAGGACAGACGCTGACGTAGAGACTGAGATTTGGCATTTGCACTTCCACCTGGGC 42

QY 995 AACACAGTAGACTCTTGTCTCAAAAAAAAAAAAAAAAAAAAAA 1033

Db 41 AACACAGAGCAAAATTCGTCTCAAAAAAAAAAAAAAAAAAAAAA 3

```

RESULT 3
LOCUS      A1733856/c
DEFINITION z191c03.y5 StrataGene colon (#937204) Homo sapiens cDNA clone
IMAGE:587332 5' similar to contains Alu repetitive element;; mRNA
sequence.
ACCESSION  A1733856
VERSION    A1733856
KEYWORDS   A1733856.1  GI:5054969
SOURCE     EST.
ORGANISM   Homo sapiens (human)
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
            1 (bases 1 to 458)
REFERENCE  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
           National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
           Tumor Gene Index
           Unpublished (1997)
COMMENT    Other ESTs: z019c03.x5
           Contact: Robert Strausberg, Ph.D.
           Email: cgapds-rt@mail.nih.gov
           This clone is available royally-free through LINT ; contact the
           IMAG Consortium (info@image.llnl.gov) for further information.
           This read is a RESEQUENCE of a previously sequenced human clone
           Original clone citation: Wash-Merck EST Project
           This read has been verified (found to hit its original self in the
           correct orientation)
           Insert Length: 1908 Std Error: 0.00
           Seq primer: -40RP from Glbcoc
           High quality sequence stop: 381.
FEATURES
             location/Qualifiers
             1..458

```

```

ORIGIN
Query Match      18.9%; Score 196.4; DB 1; Length 458
Best Local Similarity 76.9%; Pred. No. 1.2e-21;

/ld_xref="GB:4619634"
/db_xref="taxon:9606"
/clone="IMAGE:587332"
/tissue_type="tumor"
/cell_line="T84 carcinoma cell line"
/lab_host="SOLR cells (kanamycin resistant)"
/clone_1b="Stratagene colon (#937204)"
/notes="Organ: colon; Vector: pbluescript SK-
EosR; Site 2: XhoI; Cloned unidirectional
Oligo dt: T-84 colonic epithelial cell line.
insert size: 1.0 kb; Uni-ZAP XR Vector; -5' a
sequence: 5' GAATTCGACGACG 3' ~3' adaptor se
CTCGAGTTTTTTTTTTTTT 3'"

```

Matches 257; Conservative 0; Mismatches 66; Indels 11; Gaps 1;

QY 716 GAGATGAAGACTTACTTATTTAGTCCAGGGCTGGGGCCATAGCTGATGCTTGGTCC 775

Db 346 GAATTAATAATAATATCTTTTATTTTGGCCAGGGGTGGTGCATGCTGTATATCC 287

QY 776 CAGTGTCTTTCGGGGCCGAGGAGGTGAATCACTTAAGTTCAGAGTTTGGAGACCACT 835

Db 286 CAGACCTTTGGAGTCCGAGGAGGTGAATCACTTGAGTTCAGAGTTCCAGACCACT 227

QY 836 GCCCAACATGATGTAAGCTTGTCTCTACTAATAAATAAATAAGACAGGCGTGTGC 895

Db 226 GACCAACATGATGTAAGCTTGTCTCTAATAAATAAATAAGTGGCGGTGTGC 167

QY 896 ACACATCTGTAATTCAGCTACTCAGAGGCTTAACAGGAAATTCCTTGAACCTGGGA 955

Db 166 ATGGCTCTAATAATCCAGCTACTCAGGGGCTGAGGAGGAAATCCCTTGAACCTGGGA 107

QY 956 GGCAGAGTTCGAGTGA-----GCCATTGCACTCCAGCCTGGGCAACAGATGA 1004

Db 106 GGTGAGAGTTCGAGTGAAGCCGAGATCGTCCACTGCACTCCAGCTGGGTGACAGACAA 47

QY 1005 GACTCTGTCTCAAAAAAAAAAAAAAACTCGA 1038

Db 46 GACTCCATCTCAAAAAAAAAAAAAAACTCGA 13

RESULT 4
BO004374/c 621 bp mRNA linear EST 26-MAR-2002
LOCUS UI-H-E10-849-j-13-0-UI.s1 NCI CGAP_E10 Homo sapiens cDNA clone
DEFINITION IMAGE:5841132 3', mRNA sequence.

ACCESSION BO004374
VERSION BO004374.1 GI:19729274
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 621)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
The following repetitive elements were found in this cDNA
sequence: 12-301, >ALU (matched complement)
Seq primer: M13 FORWARD
POLYA=yes.

FEATURES
source Location/Qualifiers
1. 621
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5841132"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_id="NCI CGAP E10"
/note="Organ: Left Pelvis; Vector: p773-Pac (Pharmacia)
with a modified polyliner. Site 1: Ecor 1; Site 2: Not I;
NCI CGAP_E10 is a cDNA library containing the following
tissue(s): Chondrosarcoma. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double

stranded cDNA was ligated to an Ecor I adaptor, digested
with Not I, and cloned directionally into p773-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is ACACCTGCAC.
TAG TISSUE=chondrosarcoma
TAG LIB=UI-H-E10
TAG_SEQ=ACACCTGCAC"

ORIGIN

Query Match 18.9%; Score 196; DB 5; Length 621;
Best Local Similarity 81.3%; Pred. No. 1,3e-21;
Matches 244; Conservative 0; Mismatches 45; Indels 11; Gaps 1;

QY 745 GCGTGGCGCGATGCTGATGCTGTGCTCCAGTCTTTGGCGGCGGAGCGAGTGA 804

Db 301 GCGTGGGTGTGTGCTCACGCCGTGTATCTCAGCATTGTGGAGGCGCAAGCGTGTGA 242

QY 805 TCACCTTAAGGTCAGGAGTTTGAGACGACGCGCCCAATGCTGAACGTTGTCTACT 864

Db 241 TCATCTAGGTCAGGAGTTTGAGACGACGCGCCCAATGCTGAACGTTGTCTACT 182

QY 865 AAAAATCAAAAATTTAGACAGCGCGTGTGCGACATCTGTATTCAGCTACTCAGAG 924

Db 181 AAAAATCAAAAATTTAGACAGCGCGTGTGCGACATCTGTATTCAGCTACTCAGAG 122

QY 925 GCTAACACAGAAAATTCCTTGAACCTGGAGAGGAGTTGACGTGA-----G 973

Db 121 GCTGAGGACAGAGAAATCTTGGAGCTGGAGGAGGAGGTTGACGTGAGCCAGATCTGTG 62

QY 974 CCATTGACCTCCAGCCTGGGCGACACAGTGAAGTCTGTCTCAAAAAAAAAAAAAA 1033

Db 61 CCATTGACCTCCAGCCTGGGCGTATGAGTGGCTGTCTCAAAAAAAAAAAAAA 2

RESULT 5
BO017611/c 554 bp mRNA linear EST 17-JUN-2002
LOCUS UI-H-D10-auv-j-09-0-UI.s1 NCI CGAP_D10 Homo sapiens cDNA clone
DEFINITION IMAGE:5874920 3', mRNA sequence.

ACCESSION BO017611
VERSION BO017611.1 GI:19752888
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 554)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
The following repetitive elements were found in this cDNA
sequence: 1-27, >AT_richlow_complexity 133-417, >ALU (matched
complement)
Seq primer: M13 FORWARD
POLYA=yes.

FEATURES
source Location/Qualifiers
1. 554
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5874920"
/tissue_type="Lung Focal Fibrosis"

ACCESSION AQ527649.1 GI:4839803
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 456)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other GSSs: RPCI-11-348021.TU
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeef@igf.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/tcdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1..456
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7633604"
/db_xref="taxon:9606"
/clone="RPCI-11-348021"
/sex="Male"
/cell_type="lymphocytes"
/clone_1lb="RPCI-11"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC library"

ORIGIN
Query Match 18.8%; Score 195; DB 8; Length 456;
Best Local Similarity 78.0%; Pred. No. 2.1e-21;
Matches 252; Conservative 0; Mismatches 60; Indels 11; Gaps 1;
Query 721 GAAGATACTAGTTTATGTCAGAGGGCTGGCGGATGATGCTGTGTCCTGATGCCAGTG 780
DB 59 GACTTTAAATATGTTTAAATAGAGGCTGGCATGTGCTCACACTGTATATCCAGCA 118
Query 781 CTTTGGCGGGCCGAGGAGGTGATCATCTTAAGTCAGAGATTGAGACCAAGCCTGGCCA 840
DB 119 CTTTGGAGGCGAGGCGGGTGTGATCACTGAGGTCAAGATTGAGACCAAGCCTGGCCA 178
Query 841 ACATGCTGAACGTTGCTCTTACTTAAATAATCAAAAAATTAGACAGGGTGTGGTGGACACA 900
DB 179 ACATGGGAAACCCCGTCTCACATAAATAACAAAAATTAGTCAGGCGTGTGGCATGTG 238
Query 901 TCTGTAATTCAGCTACTCAGAGAGCTAACAAGAAATTCCTTGAACCTGGGAGGAG 950
DB 239 TCTGTAATTCAGCTACTCAGAGAGCTGAGAGGCTGAGAGAGAAATTGCTTGAACCTGGGAGGAG 238
Query 961 AGGTTCAGTGAAG-----CCATTGCACTCCAGCCTGGGCAACACAGTGAAGATC 1009
DB 299 AAGTTGAGTGAAGCCGAGATCGCAACCACTGCACTGAGCTGGGTAAACAAGGAGATC 358
Query 1010 TTGTTCTCAAAAAAATCAAAAAA 1032
DB 359 CTGTCTCAAAAAAATCTAAAAA 381

RESULT 8

AQ481653
LOCUS 551 bp DNA linear GSS 24-Apr-1999
DEFINITION RPCI-11-235F3.TV RPCI-11 Homo sapiens genomic clone RPCI-11-235F3, genomic survey sequence.
ACCESSION AQ481653
VERSION AQ481653.1 GI:4669057
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 551)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other GSSs: RPCI-11-235F3.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeef@igf.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/tcdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1..551
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7589978"
/db_xref="taxon:9606"
/clone="RPCI-11-235F3"
/sex="Male"
/cell_type="lymphocytes"
/clone_1lb="RPCI-11"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC library"

ORIGIN
Query Match 18.8%; Score 194.8; DB 8; Length 551;
Best Local Similarity 79.5%; Pred. No. 2.1e-21;
Matches 229; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
Query 745 GGCTGGGCGGATAGCTGATGCTGTGTCCTTGGCGGGCCGAGGAGGTGA 804
DB 247 GGCTGGGCGGAGTGGCCAGCTGTATATCCAGCACTTGGGAGGCTGAGATGAGTGA 306
Query 805 TCACTTAAGTCAAGAGTTTGAAGACCAAGCTGGCCAAAGTGAACCTGTCTACT 864
DB 307 TCACTGAGGTCAAGCAATTTGAAGACCAAGCTGGCCAAAGTGAACCTGTCTACT 366
Query 865 AAAAATAAATAAATTAGACAGAGCGTGTGGACACATCTGTAATTCAGCTACTCAGAG 924
DB 367 AAAAATAAATAAATTAGCTGGGCGTGTGGTGGCAAGCTCTTGTAGTCCAGCTACCGGGG 426
Query 925 GCTTAACACAGAAATTTCTTGAACCTGGAGGCAAGAGTTGACAGTGAAGCAATTGCACTC 984
DB 427 GCTGAAGAGGAATTTGCTTGAACCCGGGAGGACAGGTTGCAGTGAAGCACCTGCACTC 486
Query 985 CAGCCTGGGCAACAGTGAAGCTTGTCTCAAAAAAATCAAAAAA 1032
DB 487 CAGTCTGGGTGATGAGTGAAGTCTGTCTCAAAAAAATCAAAAAA 534

RESULT 9

FEATURES
This clone (DREF26860380) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers

Query Match	18.8%	Score 194.8	DB 5	Length 681
Best Local Similarity	78.4%	Pred. No. 2e-21		
Matches 250	Conservative 0	Mismatches 58	Indels 11	Gaps 1
QY	731	AGTTTAACTCAGGGGCTGGGGCGCATGA	CTGATCCTGTGTCCTCAGTCTTTGGGGG	790
Db	418	AATATGTGTATAGCTCTGGGTGGGTGGCT	CATGCTGTAAATCCAGACATTGGAGG	359
QY	791	CCGAGGCGAGTGGATCATCTTAAAGT	CAGAGATTGGAGACCAAGCTTGCCCAACAGTGTGA	850
Db	358	CCAAAGCGAGTGGATMACTGAGGTCA	GGATTGAAACAGCCTGAACCAACATGGTGA	299
QY	851	ACGTTGCTCTACTAAATAACA	AAATTGACAGGGGTGTGGCACATCTGTAAATTC	910
Db	298	ATCCAGTCTCTAATTAATACAAA	ATTAGCCAGGGGTGGTCGTGCAATTTGTAAATCC	239
QY	911	CAGTACTCAGAGGCTAAACA	GAGAAATTTCTTGTGAACCTGGAGGACAGAGTTGCAAT	970
Db	238	CAGTACACAGGAGGCTAGGCA	GAGAAATTTGCTTGAACCTGGGGGGCGGAGTTTGCAAT	179
QY	971	GA-----GCCATTGCACTCC	AGCCTGGGCAACACAGTGAACCTTTGTCTCAA	1019
Db	178	GAGCCAGGTGCGCGCATTTG	CACTCCAGCTGGGGCGACAGATGAACATCTTCCAAA	119
QY	1020	AAAAAAAAAAAAAACTGA	1038	
Db	118	AAAAAAAAAAAACAAAAGA	100	

FEATURES

Ingolstaedter **lanstr.1**, D-85764 Neuberg, Germany
This is the 5' sequence of the clone insert. Clone from S. Wiemann,
Molecular Genome Analysis, German Cancer Research Center (DKFZ),
Email s.wiemann@dkfz-heidelberg.de; mforschung GmbH in Berlin,
Germany. Please contact RZPD for ordering:
<http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp45912062>
Further information about the clone and the sequencing project is
available at <http://mips.gsf.de/projects/cdna/>.
Location/Qualifiers

Query Match	18.7%	Score 194.2	DB 7	Length 452
Best Local Similarity	78.4%	Pred. No. 2.8e-21		
Matches 250	Conservative	0	Mismatches 58	Indels 11
				Gaps 1
QY	719	ATGAAGATACCTAGTTTGTAGTCCAGGGGCTGGGCGCATAGCTGATGCTCGTGTGATCCAG		778
Db	374	ATGAGTATATTAAATTAAAGAGGGGCAAGGCTGGGCACTGATGTTCAAGCCCTGTAATCCAG		315
QY	779	TGCTTTGGGGGCGGAGGCGAGGTGGATCACTTAAAGTCAGAGGTTTGAGACCAAGCTGGCC		838
Db	314	CATTTTGGGAGGCCAAGGCAAGGTGGATCAACTGAGGTCAGAGATTGTGAGACCAAGCTGAC		255
QY	839	CAACATGGTGAACGGTTGCTCTTACTTAAAAATTCAAAAATTAAACAGCGCTGTGGCAACA		898
Db	234	CACATGGTGAACCCCGTCTCTTACTTAAAAATTCAAAAATTAAACAGGTAATGGTGGCGCC		195
QY	899	CATCTGTAAATTCAGACTACTCAGAGGCTAACAAGAAAAATTCCTTGAACCTGGAGAGCC		958
Db	194	TGCCGTGTAATCCAGGCTAATCTCAGGAGGCTGAGGCAAGAGATCACTTGAACCCGGAGGCC		135
QY	959	AGAGGTTGCACTG-----GCCATTGCACTCCAGGCTGGGCAACACAGTGAGAC		1007
Db	134	AGAGGTTGCACTGAGCCAAATGCAAGCCATTGCACTCCAGCGTGGCAACAGAGTGAGCC		75
QY	1008	TCTTGTCTCAAAAAAAAAA	1026	
Db	74	TCTGACAAAAAAAAA	56	

RESULT 11					
BM667682/c					
LOCUS	BM667682	504 bp	mRNA	linear	EST 27-FEB-2002
DEFINITION	UI-E-DX0-ags-f-14-0-UI.s1 UI-E-DX0 Homo sapiens cDNA clone UI-E-DX0-ags-f-14-0-UI 3', mRNA sequence.				

ACCESSION	BM667682
VERSION	BM667682.1
KEYWORDS	GI:1897513
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	1 (bases 1 to 504)
JOURNAL	Bonaldo,M.F., Lennon,G. and Soares,M.B.
MEDLINE	Normalization and subtraction: two approaches to facilitate gene discovery
PUBMED	Genome Res. 6 (9), 791-806 (1996)
COMMENT	97044477
	8889548
	Contact: Soares, MB

Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.researchgen.com).
The following repetitive elements were found in this cDNA
sequence: 1-27, >At_richlow_complexity 132-416, >MLU (matched
compliment)
Seq primer: M13 Forward
POLYA=Yes.

FEATURES

Location/Qualifiers

```

1. .504
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-DX0-aga-F-14-0-UI"
/tissue_type="fetal eyes"
/dev_stage="fetal"
/lab_host="DHI08 (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-DX0"
/notes="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-DX0 is a cDNA library containing the following
tissue(s): fetal eyes. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is AGAATCAGA. This library
was created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI).
TAG_TISSUE=human fetal eyes
TAG_LIB=UI-E-DX0
TAG_SEQ=AGAATCAGA"

```

ORIGIN

Query Match	18.7%	Score 194.2	DB 4	Length 504
Boot Local Similarity	78.4%	Pred. No. 2.7e-21		
Matches 250	Conservative 0	Mismatches 55	Indels 11	Gaps 1
Qy	731	AGTTTATGTCAGAGGGCTGGGCGCATAGCTGATGCTCTGTGTGCCAGTCTTTGCGGGG		790
Dy	430	AATATGTTGTATAGGCTGGGCGCGTGCATAGCTGTATCCACACATTTGGGAGG		371
Qy	791	CCGAGGCGAGTGATCATCTTAAGTACAGAGATTTGAGACCAAGCCTGCCCAACATGTGAA		850
Dy	370	CCAAGGCGAGTGATCATCTTAAGTACAGAGATTTGAGACCAAGCCTGCCCAACATGTGTA		311

Qy 851 ACCTTGTCTCACTAAATAATCAAAATAATTAACAGCGGTGGTGCACACATCTGTAATTC 910

Db 310 ATCCAGTCTCACTAAATAATCAAAATAATTAACAGCGGTGGTGCATTTGTAAATCC 251

Qy 911 CAGCTACTCAGGAGGCTTAACACAGAGAAATTCCTTGAACCTGGAGCAGAGGTTGCAGT 970

Db 250 CAGCTACACAGGAGGCTGAAGCAGAGAAATTTGCTTGAACCTGGGGGGGCGAGGTTGCANT 191

Qy 971 GA-----GCCATTGCATCTCCAGCCTGGGCAACACAGTGAGACTTGTCTCAAA 1015

Db 190 GAGCCAAAGTGCGCGCATTTGCATCTCCAGCCTGGGCGACAGAGTGACACTTCTCAAAAAAAG 131

Qy 1020 AAAAAAAAAAAACTGA 1038

Db 130 AAAAAAACAATAAAGA 112

RESULT	12
AIO73373/c	
LOCUS	AIO73373
DEFINITION	515 bp mRNA linear EST 06-AUG-1998 o033a11.x1 Searaea NSF_P8_9W_OT_PA.P_Sl Homo sapiens cDNA clone IMAGE156604 J similar to contains Alu repetitive element/, mRNA sequence.
ACCESSION	AIO73373
VERSION	AIO73373.1 GI:3400017
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE
AUTHORS
TITLE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 515)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabbs@emall.nih.gov
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.lln.gov) for further information.
Seq primer: -40m13 fwd, RT from Amersham
High quality sequence stop: 454.

FEATURES

```

/organism="Homo sapiens"
/mol_type="mrna"
/ds_xref="taxon:9606"
/clone="IMAGE:1566044"
/lab_host="DH10B"
/clone_lib="Scorers NSF Pg. 9W OT PA. p.S1"
/notes="Organ: pooled; Vector: pTfT3-Pac (pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and sa circles were made in vitro.
Following RNP hybridization, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and clonesId: Scorers NBHSF pool 1:
309384-310919, 323208-325895 Scorers NB2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Scorers NB2HF8-9W pool 1:
758280-760583, 772014-774407 Scorers NBHPA pool 1:
304776-306311, 320136-322823, 326280-326663 Scorers NBHOT
pool 1: 723270-726407, 739080-740939 Subtraction by Bencho
Scorers and M. Fatima Donaldso."

```

ORIGIN

Query Match 18.7%; Score 194.2; DB 1; Length 515;
Best Local Similarity 78.4%; Pred. No. 2,7e-21;
Matches 250; Conservative 0; Mismatches 56; Indels 11; Gaps 1

```

Db      423 AATATGTGTATAGTGGTGGTGGCTGCGGTGCTATCCCTGTAATCCAGACCTTTGGAGG 364
QY      791 CCGAGGCGAGGTGATGATCTTAAGGTGAGAGTTTGAGACCGAGCTCCCAACATGTGTGA 850
Db      363 CCAAGGACAGGTGATGATCCTGAGGTGAGAGTTTGAGACCGAGCTCCCAACATGTGTGA 304
QY      851 ACCTTGTCTTACTTAAATAATGAAAAATGAGACGCGGTGTGGGACACATCTGTATTC 910
Db      303 ATCAAGTCTTACTTAAATAATGAAAAATGAGACGCGGTGTGGGACACATCTGTATTC 244
QY      911 CAGCTACTGAGAGGCTTAACACAGAAAAATTCCTTGAACCTGGGAGGAGGAGTTGCACT 970
Db      243 CAGCTACACAGAGGCTGAGAGGAGAAATGCTTGAACCTGGGAGGAGGAGTTGCACT 184
QY      971 GA-----GCCATTGCACTCCAGCCTGGGCAACAGCAGTGAAGCTTTGTCTCAA 1019
Db      183 GAGCCAGGTGCGGCGCATTTGCACTCCAGCCTGGGCAACAGTGAAGCTTTCAAAAAAC 124
QY      1020 AAAAAAAAAAAAAAATCGA 1038
Db      123 AAAAAACAAAAACAAAAAGA 105

```

```

RESULT 13
LOCUS   AQJ14507/c 524 bp DNA linear GSS 04-MAY-1999
DEFINITION RPI11-103B6.TV RPI1-11 Homo sapiens genomic clone RPI1-11-103B6,
ACCESSION AQJ14507
VERSION   AQJ14507.1 GI:4045970
KEYWORDS GSS.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 524)
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,B., Wible,C., de Jong,P. and Venter,J.C.
Use of human BAC End Sequences for Sequence-Ready Map Building
Unpublished (1998)
Other GSS: RPI11-103B6.TI
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPI1-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: T7
Class: BAC ends.

```

FEATURES

source Location/Qualifiers

1..524

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="GDB:7539197"

/db_xref="taxon:9606"

/clone="RPI1-11-103B6"

/sex="Male"

/cell_type="Lymphocytes"

/clone_lib="RPI1-11"

/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPI11 Human Male BAC library"

ORIGIN

Query Match 18.7%; Score 194.2; DB 8; Length 524;

Best Local Similarity 79.8%; Pred. No. 2.7e-21;

```

Matches 229; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY      743 GGGCTGGGCGCGATAGCTGATGCTGAGGCCAGGCTTTGGGGCCGAGCGAGTGT 802
Db      415 GAGCTGGGCGCGGTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 356
QY      803 GATCACTTAAGTGAAGAGTTTGAAGACCAAGCTGCGCCCAACATGATGTAACGTTCTCTA 862
Db      355 GATCACTTAAGTGAAGAGTTTGAAGACCAAGCTGCGCCCAACATGATGTAACGTTCTCTA 296
QY      863 CTAATAATATCAAAATTAAGACGCGGTGTGGGACACATCTGTATTCAGTCTCAG 922
Db      295 CTAATAATATCAAAATTAAGACGCGGTGTGGGACACATCTGTATTCAGTCTCAG 236
QY      923 AGGCTAACACAGAAAAATTCCTTGAACCTGGGAGGAGAGTGTGAGTGAAGCATTGAC 982
Db      235 AGGCTAACGCGGAGAAATGCTTGAACCTGAAGAGGAGAAAGTTGCACTGAACACCGCAC 176
QY      983 TCCAGCCTGGGCAACACAGTGAAGCTTGTCTCAAAAAAAAAAAAAA 1029
Db      175 TCCAGCCTGGGTGACGAGGAGAGTCTATCTCAAAAAAAAAAAAAA 129

```

```

RESULT 14
LOCUS   BU626395 724 bp mRNA linear EST 23-SEP-2002
DEFINITION BU626395 BU626395
ACCESSION BU626395
VERSION   BU626395.1 GI:23292610
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 724)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@iowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-27, >AT rich#low complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=yes.

```

FEATURES

source

Location/Qualifiers

1..724

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-H-DF0-beu-o-10-0-UI"

/tissue_type="Subchondral Bone"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NCI CGAP DF0"

/note="Organ: Bone; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site_1: Sca I; Site_2: Not I; NCI CGAP DF0 is a cDNA library containing the following tissue(s): Subchondral Bone. The library was constructed according to Bonaldi, Lemon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Sca I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of

first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GTTAAGCTC.
TAG_TISSUE=subchondral bone
TAG_LIB=UI-H-DFO
TAG_SEQ=GTTAAGCTC"

ORIGIN

```
Query Match      18.7%; Score 194.2; DB 5; Length 724;
Best Local Similarity 78.4%; Pred. No. 2.4e-21;
Matches 250; Conservative 0; Mismatches 58; Indels 11; Gaps 1;

Qy 731 AGTTTATGTCAGAGGCGTGGCGCATATGATGCTGTGTCCTTGGCGGG 790
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 431 AATATGCTGTATAGCTGCGTGGTGGCTCATGCTCTTAATCCACACATTTGGAGG 372

Qy 791 CCGAGGAGGTGATCATCTTAAGGTCAGAGTTTGAACACAGCTGCCCAATGTGAA 850
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 371 CCAAGGAGGTGATCATCTGAGGTCAGAGTTTGAACACAGCTGCACATGTGAA 312

Qy 851 AGCTGTCTCTAATAAATCAAAATTAACACAGGCGTGTGGCACACATCTTAATTC 910
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 311 ATCCAGCTCTAATAAATCAAAATTAACACAGGCGTGTGGCATTTGTAAATCC 252

Qy 911 CAGCTACTCAGAGAGCTTAACAGAGAAATTCCTTGAACCTGGAGGAGAGTTGCAGT 970
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 251 CAGCTTACACAGAGAGCTGAGGACAGAAATTCCTTGAACCTGGAGGAGAGTTGCAGT 192

Qy 971 GA-----GCCATTGCACTCCAGCTGGGCAACACAGTAGAGCTCTTGTCTCAA 1019
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 191 GAGCCAGATGTCGCCCATTTGCACTCCAGCTGGGCGAGAGTAGACACTTCTCAAAAAC 132

Qy 1020 AAAAAAAAAAAAAAAGCTGA 1038
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 131 AAAAAAACAACAAAAAGA 113
```

RESULT 15

LOCUS AO415039 540 bp DNA linear GSS 23-MAR-1999
DEFINITION RPCI-11-203B9.TV RPCI-11 Homo sapiens genomic clone RPCI-11-203B9,
genomic survey sequence.

ACCESSION AO415039
VERSION AO415039.1 GI:4474008
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 540)
AUTHORS Zhao,S., Adams,W.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
MAP Building
JOURNAL Unpublished (1997)
COMMENT Other GSSs: RPCI-11-203B9.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@ligr.org

Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genet cs (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1..540

FEATURES
SOURCE

```
/organism="Homo sapiens"  
/mol_type="genomic DNA"  
/db_xref="GDB:7577672"  
/db_xref="taxon:9606"  
/clone="RPCI-11-203B9"  
/sex="Male"  
/cell_type="Lymphocytes"  
/clone_lib="RPCI-11"  
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;  
RPCI11 Human Male BAC Library"
```

ORIGIN

```
Query Match      18.7%; Score 193.6; DB 8; Length 540;
Best Local Similarity 80.3%; Pred. No. 3.3e-21;
Matches 244; Conservative 0; Mismatches 49; Indels 11; Gaps 1;

Qy 743 GGGGCTGGCGCGATGACGTATGCTGTGCTCCAGTGCTTTGGCGGGCCGAGGCAAGTG 802
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 179 GGGGCGAGGTGTGTACTGATGCTGTAAACCCAGACCTTTGGAGGCGAGGCAAGTG 238

Qy 803 GATCACTTAAGGTAGAGAGTTTGAACACAGCTGCTGCCCAACATGTGAAAGCTTGTCTTA 862
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 239 GATCACTTCAAGCTCAGAGAGTTTGAACACAGCTGCTGCCCAACATGTGAAAGCTTGTCTTA 298

Qy 863 CTAAAAATACAAAATTAAGACAGGCGTGTGGGACACATCTGTAAATCCAGCTACTCAGG 922
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 299 CTAAAAATACAAAATTAAGTACAGCGTGTGGGACACATCTGTAAATCCAGCTACTCAGG 358

Qy 923 AGGCTAACACAGGAAAATTCCTTGAACCTGGAGGCAAGGTTGCAAGTGA----- 972
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 359 AGGCTGAGGCGAGGAATCACTTGAACCTGGAGGCAAGGTTGCAAGTGAAGGCCAGATCG 418

Qy 973 -GCCATTGCACTCCAGCTGGGCAACACAGTAGAATCTTCTCAAAAAAAAAAAAAA 1031
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 419 TGCCATTGCACTCCAGCTGGAGGCAAGTAGAATCTCAAAAAAAAAATAATTA 478

Qy 1032 AACT 1035
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 479 TAAT 482
```

Search completed: March 6, 2005, 06:38:16
Job time : 3348 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 6, 2005, 07:53:15 ; Search time 59 Seconds
(without alignments)
61.997 Million cell updates/sec

Title: US-10-664-358-549

Perfect score: 260

Sequence: 1 MNLIGMIFSMCGMLKLKMKC.....ISFANSRSSEDTKOMSSFM 49

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCUTS.COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	260	100.0	49	4	US-09-482-273-180 Sequence 180, App
2	260	100.0	70	3	US-09-188-930-313 Sequence 313, App
3	260	100.0	70	4	US-09-312-283C-313 Sequence 313, App
4	260	100.0	71	4	US-09-482-273-116 Sequence 116, App
5	260	100.0	106	3	US-09-247-155-90 Sequence 90, App1
6	260	100.0	106	4	US-09-513-999C-8115 Sequence 8115, App
7	217	83.5	114	4	US-09-270-767-33341 Sequence 33341, A
8	217	83.5	114	4	US-09-270-767-48558 Sequence 48558, A
9	61	23.5	240	4	US-09-270-767-39637 Sequence 39637, A
10	61	23.5	240	4	US-09-270-767-54854 Sequence 54854, A
11	58.5	22.5	322	4	US-09-254-227A-5 Sequence 5, App11
12	54.5	21.0	322	4	US-09-254-227A-3 Sequence 3, App11
13	53	20.4	223	4	US-09-465-901-16 Sequence 16, App1
14	52	20.0	65	4	US-09-248-796A-25087 Sequence 25087, A
15	52	20.0	190	4	US-09-248-796A-15809 Sequence 15809, A
16	51	19.6	223	4	US-09-465-901-14 Sequence 14, App1
17	51	19.6	223	4	US-09-465-901-32 Sequence 32, App1
18	50.5	19.4	316	4	US-09-252-991A-17312 Sequence 17312, A
19	50.5	19.4	352	4	US-09-134-000C-5338 Sequence 5338, App
20	50	19.2	139	4	US-09-270-767-37702 Sequence 37702, A
21	50	19.2	139	4	US-09-270-767-52919 Sequence 52919, A
22	50	19.2	239	4	US-09-489-039A-10233 Sequence 10233, A
23	50	19.2	848	4	US-09-543-681A-7615 Sequence 7615, App
24	49.5	19.0	202	4	US-09-252-991A-19646 Sequence 19646, A
25	49.5	19.0	585	4	US-09-248-796A-20096 Sequence 20096, A
26	49.5	19.0	741	4	US-09-436-699C-22 Sequence 22, App1
27	49	18.8	54	4	US-09-621-976-5390 Sequence 5390, App

28	49	18.8	118	4	US-09-270-767-62231 Sequence 62231, A
29	49	18.8	154	4	US-09-248-796A-20059 Sequence 20059, A
30	49	18.8	223	4	US-09-465-901-36 Sequence 36, App1
31	49	18.8	348	4	US-09-270-767-46628 Sequence 46628, A
32	49	18.8	373	4	US-09-248-796A-17973 Sequence 17973, A
33	49	18.8	462	3	US-09-165-241-1 Sequence 1, App11
34	48	18.5	54	4	US-09-621-976-5389 Sequence 5389, App
35	48	18.5	125	4	US-08-469-260A-412 Sequence 412, App
36	48	18.5	125	4	US-08-468-446-412 Sequence 412, App
37	48	18.5	125	4	US-08-467-344A-412 Sequence 412, App
38	48	18.5	125	4	US-08-424-550B-412 Sequence 412, App
39	48	18.5	213	2	US-08-808-550-35 Sequence 35, App1
40	48	18.5	223	4	US-09-465-901-42 Sequence 42, App1
41	48	18.5	307	4	US-09-902-540-15716 Sequence 15716, A
42	48	18.5	315	4	US-09-543-681A-7121 Sequence 7121, App
43	48	18.5	367	4	US-09-236-840A-3 Sequence 3, App11
44	48	18.5	376	4	US-09-721-870-113 Sequence 113, App
45	47.5	18.3	106	4	US-09-252-991A-29961 Sequence 29961, A

ALIGNMENTS

```
RESULT 1
US-09-482-273-180
; Sequence 180, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; EARLIER FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 180
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-482-273-180

Query Match      100.0%; Score 260; DB 4; Length 49;
Best local similarity 100.0%; Pred. No. 6.8e-29;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1 MNLIGMIFSMCGMLKLKMKCAWAVVCSPTISFANSRSSEDTKOMSSFM 49
Db      1 MNLIGMIFSMCGMLKLKMKCAWAVVCSPTISFANSRSSEDTKOMSSFM 49

RESULT 2
US-09-188-930-313
; Sequence 313, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Muriel, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
```

```
/ NUMBER OF SEQ ID NOS: 348
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 313
/ LENGTH: 70
/ TYPE: PRT
/ ORGANISM: Human
US-09-188-930-313
```

```
Query Match
Best Local Similarity 100.0%; Score 260; DB 3; Length 70;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MNILGIFSMCGIMLKLKWCAMVAVYCSFISFANSRSSSEDTKOMSSFM 49
DB 1 MNILGIFSMCGIMLKLKWCAMVAVYCSFISFANSRSSSEDTKOMSSFM 49
```

```
RESULT 3
US-09-312-283C-313
/ Sequence 313, Application US/09312283C
/ Patent No. 6573095
/ GENERAL INFORMATION:
/ APPLICANT: Watson, James D.
/ APPLICANT: Strachan, Lorna
/ APPLICANT: Sleeman, Matthew
/ APPLICANT: Onrust, Rene
/ APPLICANT: Murlson, James G.
/ APPLICANT: Kumble, Krishanand D.
/ TITLE OF INVENTION: Compositions Isolated from Skin Cells
/ FILE REFERENCE: 11000.1011c2
/ CURRENT APPLICATION NUMBER: US/09/312,283C
/ NUMBER OF SEQ ID NOS: 425
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 313
/ LENGTH: 70
/ TYPE: PRT
/ ORGANISM: Mouse
US-09-312-283C-313
```

```
Query Match
Best Local Similarity 100.0%; Score 260; DB 4; Length 70;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MNILGIFSMCGIMLKLKWCAMVAVYCSFISFANSRSSSEDTKOMSSFM 49
DB 1 MNILGIFSMCGIMLKLKWCAMVAVYCSFISFANSRSSSEDTKOMSSFM 49
```

```
RESULT 4
US-09-482-273-116
/ Sequence 116, Application US/09482273
/ Patent No. 6534631
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: 71 Human Secreted Proteins
/ FILE REFERENCE: P2030P1
/ CURRENT APPLICATION NUMBER: US/09/482,273
/ EARLIER FILING DATE: 2000-01-13
/ EARLIER APPLICATION NUMBER: PCT/US99/15849
/ EARLIER FILING DATE: 1999-07-14
/ EARLIER APPLICATION NUMBER: 60/092,921
/ EARLIER FILING DATE: 1998-07-15
/ EARLIER APPLICATION NUMBER: 60/092,922
/ EARLIER FILING DATE: 1998-07-15
/ EARLIER APPLICATION NUMBER: 60/092,956
/ EARLIER FILING DATE: 1998-07-15
/ NUMBER OF SEQ ID NOS: 267
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 116
/ LENGTH: 71
/ TYPE: PRT
```

```
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (71)
/ OTHER INFORMATION: Xaa equals stop translation
US-09-482-273-116
```

```
Query Match
Best Local Similarity 100.0%; Score 260; DB 4; Length 71;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MNILGIFSMCGIMLKLKWCAMVAVYCSFISFANSRSSSEDTKOMSSFM 49
DB 1 MNILGIFSMCGIMLKLKWCAMVAVYCSFISFANSRSSSEDTKOMSSFM 49
```

```
RESULT 5
US-09-247-155-90
/ Sequence 90, Application US/09247155A
/ Patent No. 6312822
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Milne Edwards, Jean-Baptiste
/ APPLICANT: Duclert, Aymeric
/ APPLICANT: Bouqueloret, Lydie
/ TITLE OF INVENTION: Complementary DNAs
/ FILE REFERENCE: GENSET.021A
/ CURRENT APPLICATION NUMBER: US/09/247,155A
/ EARLIER FILING DATE: 1999-02-09
/ EARLIER APPLICATION NUMBER: 60/074,121
/ EARLIER FILING DATE: 1998-02-09
/ EARLIER APPLICATION NUMBER: 60/081,563
/ EARLIER FILING DATE: 1998-04-13
/ EARLIER APPLICATION NUMBER: 60/096,116
/ EARLIER FILING DATE: 1998-08-10
/ EARLIER APPLICATION NUMBER: 60/099,273
/ EARLIER FILING DATE: 1998-10-04
/ NUMBER OF SEQ ID NOS: 182
/ SOFTWARE: Patent.pm
/ SEQ ID NO 90
/ LENGTH: 106
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SIGNAL
/ LOCATION: -71...-1
US-09-247-155-90
```

```
Query Match
Best Local Similarity 100.0%; Score 260; DB 3; Length 106;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MNILGIFSMCGIMLKLKWCAMVAVYCSFISFANSRSSSEDTKOMSSFM 49
DB 37 MNILGIFSMCGIMLKLKWCAMVAVYCSFISFANSRSSSEDTKOMSSFM 85
```

```
RESULT 6
US-09-513-999C-8115
/ Sequence 8115, Application US/09513999C
/ Patent No. 6783961
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Milne Edwards, J.B.
/ APPLICANT: Duclert, A.
/ APPLICANT: Giordano, J.Y.
/ TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
/ FILE REFERENCE: 59, US2,REG
/ CURRENT APPLICATION NUMBER: US/09/513,999C
/ EARLIER FILING DATE: 2000-02-24
/ PRIOR FILING DATE: 1999-02-26
/ NUMBER OF SEQ ID NOS: 36681
/ SOFTWARE: Patent.pm
```

SEQ ID NO 8115
LENGTH: 106
TYPE: PRT
ORGANISM: Homo sapiens
US-09-513-899C-8115

Query Match 100.0%; Score 260; DB 4; Length 106;
Best Local Similarity 100.0%; Pred. No. 1.6e-28;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMCGMLKXKMCAMVAVYCSFISFANSRSSEDTKQWMSFM 49
DB 37 MNLIGMIFSMCGMLKXKMCAMVAVYCSFISFANSRSSEDTKQWMSFM 85

RESULT 7

US-09-270-767-33341
Sequence 33341, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 33341
LENGTH: 114
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-33341

Query Match 83.5%; Score 217; DB 4; Length 114;
Best Local Similarity 77.6%; Pred. No. 1.6e-22;
Matches 38; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMCGMLKXKMCAMVAVYCSFISFANSRSSEDTKQWMSFM 49
DB 43 MNLIGMIFSMCGMLKXKMCAMFALYCSCISFANSRSASDAKQVLSFM 91

RESULT 8

US-09-270-767-48558
Sequence 48558, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 48558
LENGTH: 114
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-48558

Query Match 83.5%; Score 217; DB 4; Length 114;
Best Local Similarity 77.6%; Pred. No. 1.6e-22;
Matches 38; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMCGMLKXKMCAMVAVYCSFISFANSRSSEDTKQWMSFM 49
DB 43 MNLIGMIFSMCGMLKXKMCAMFALYCSCISFANSRSASDAKQVLSFM 91

RESULT 9
US-09-270-767-39637
Sequence 39637, Application US/09270767
Patent No. 6703491

GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 39637
LENGTH: 240
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-39637

Query Match 23.5%; Score 61; DB 4; Length 240;
Best Local Similarity 24.1%; Pred. No. 1.4;
Matches 13; Conservative 10; Mismatches 21; Indels 10; Gaps 1;

QY 2 NLLGMIFSMCGMLKXKMCAMVAVYCSFISF-----ANSRSSEDTKQW 45
DB 38 NLYCTVFAMCGXKTKIKNGCLIVLYCLVFFFXAIWQDKYXNENSDSYEQFL 91

RESULT 10

US-09-270-767-54854
Sequence 54854, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 54854
LENGTH: 240
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-54854

Query Match 23.5%; Score 61; DB 4; Length 240;
Best Local Similarity 24.1%; Pred. No. 1.4;
Matches 13; Conservative 10; Mismatches 21; Indels 10; Gaps 1;

QY 2 NLLGMIFSMCGMLKXKMCAMVAVYCSFISF-----ANSRSSEDTKQW 45
DB 38 NLYCTVFAMCGXKTKIKNGCLIVLYCLVFFFXAIWQDKYXNENSDSYEQFL 91

RESULT 11

US-09-254-227A-5
Sequence 5, Application US/09254227A
Patent No. 6696257
GENERAL INFORMATION:
APPLICANT: Ahmad, Sultan
APPLICANT: Banville, Denis
APPLICANT: Portin, Yves
APPLICANT: Lembo, Paola
APPLICANT: O'Donnell, Dajan
APPLICANT: Shi-Hsiang, Shen
TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
FILE REFERENCE: 81823/268117
CURRENT APPLICATION NUMBER: US/09/254,227A
CURRENT FILING DATE: 1999-03-03
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patentin version 3.0
SEQ ID NO 5
LENGTH: 322

TYPE: PRT
ORGANISM: Homo sapiens
US-09-254-227A-5

Query Match 22.5%; Score 58.5; DB 4; Length 322;
Best Local Similarity 25.0%; Pred. No. 4.2;
Matches 14; Conservative 10; Mismatches 19; Indels 13; Gaps 2;

QY 6 MIFMCGMLMKWCA-----WVAYCS-----FISFANSRSSSDTKQMMSSF 48
DB 223 LVFLGCLPFGIGWALFSRIHLDKVLFCHVHLVSIPLSALNNSANPLIYFFVGSF 278

RESULT 12
US-09-254-227A-3
Sequence 3, Application US/09254227A
Patent No. 6696257
GENERAL INFORMATION:
APPLICANT: Ahmad, Sultan
APPLICANT: Banville, Denis
APPLICANT: Fortin, Yves
APPLICANT: Lembo, Paola
APPLICANT: O'Donnell, Dajan
APPLICANT: Shi-Heiang, Shen
TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
FILE REFERENCE: 81823/268117
CURRENT APPLICATION NUMBER: US/09/254,227A
CURRENT FILING DATE: 1999-03-03
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patentin version 3.0
SEQ ID NO 3
LENGTH: 322
TYPE: PRT
ORGANISM: Homo sapiens
US-09-254-227A-3

Query Match 21.0%; Score 54.5; DB 4; Length 322;
Best Local Similarity 23.2%; Pred. No. 15;
Matches 13; Conservative 11; Mismatches 19; Indels 13; Gaps 2;

QY 6 MIFMCGMLMKWCA-----WVAYCS-----FISFANSRSSSDTKQMMSSF 48
DB 223 LVFLGCLPFGIGWALFSRIHLDKVLFCHVHLVSIPLSALNNSANPLIYFFVGSF 278

RESULT 13
US-09-465-901-16
Sequence 16, Application US/09465901
Patent No. 6492143
GENERAL INFORMATION:
APPLICANT: Reed, Randall
APPLICANT: Yau, King-Wai
APPLICANT: Krautwurst, Dietmar
TITLE OF INVENTION: Olfactory Receptor Expression Libraries
TITLE OF INVENTION: ad Methods of Making and Using Them
FILE REFERENCE: 001107.00105
CURRENT APPLICATION NUMBER: US/09/465,901
CURRENT FILING DATE: 2001-12-17
PRIOR APPLICATION NUMBER: 60/112,605
PRIOR FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 57
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 223
TYPE: PRT
ORGANISM: Mus musculus
FEATURE:
OTHER INFORMATION: PCR primer
US-09-465-901-16

Query Match 20.4%; Score 53; DB 4; Length 223;
Best Local Similarity 31.8%; Pred. No. 16;
Matches 14; Conservative 12; Mismatches 16; Indels 2; Gaps 1;

QY 8 FSWGMLMKWCAWVAYCSFISFANSRSSSDTK--QMMSSFM 49
DB 175 FSTCSSHLIVSVSYSSVFCAYVSPASSYSPEKSKVTSLVYBFL 218

RESULT 14
US-09-248-796A-25087
Sequence 25087, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 25087
LENGTH: 65
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-25087

Query Match 20.0%; Score 52; DB 4; Length 65;
Best Local Similarity 35.3%; Pred. No. 5.5;
Matches 12; Conservative 4; Mismatches 18; Indels 0; Gaps 0;

QY 13 LMLKMKWCAWVAYCSFISFANSRSSSDTKQMS 46
DB 6 LKDLKMGKTKCNVCRFPFKKQKTKTKTKTS 39

RESULT 15
US-09-248-796A-15809
Sequence 15809, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 15809
LENGTH: 190
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-15809

Query Match 20.0%; Score 52; DB 4; Length 190;
Best Local Similarity 28.6%; Pred. No. 18;
Matches 12; Conservative 9; Mismatches 13; Indels 8; Gaps 1;

QY 5 GMIFMCGMLMKWCAWVAYCSFISFANSRSSSDTKQMS 46
DB 132 GILFTCC-----AOFVNLICLFVYISNRGSSRDLSKIPS 165

Search completed: March 6, 2005, 08:10:26
Job time : 60 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 6, 2005, 08:08:21 ; Search time 93 Seconds

(without alignments)
173.368 Million cell updates/sec

Title: US-10-664-358-549

Perfect score: 260
Sequence: 1 MNLLGTMFSCGLMKLKWC.....ISFANSSSDTQMMSSFM 49Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1391452 seqs, 329044822 residues

Total number of hits satisfying chosen parameters: 1391452

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/prodata/2/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/prodata/2/pubppaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/prodata/2/pubppaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/prodata/2/pubppaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/prodata/2/pubppaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/prodata/2/pubppaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/prodata/2/pubppaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/prodata/2/pubppaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/prodata/2/pubppaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/prodata/2/pubppaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/prodata/2/pubppaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/prodata/2/pubppaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/prodata/2/pubppaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/prodata/2/pubppaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/prodata/2/pubppaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/prodata/2/pubppaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/prodata/2/pubppaa/US10E_NEW_PUB.pep:*
- 18: /cgn2_6/prodata/2/pubppaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/prodata/2/pubppaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/prodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	260	100.0	49	9	US-09-984-245-187
2	260	100.0	49	10	US-09-984-271-180
3	260	100.0	49	10	US-09-984-271-180
4	260	100.0	49	10	US-09-966-262-187
5	260	100.0	49	10	US-09-983-966-187
6	260	100.0	49	14	US-10-059-395-187
7	260	100.0	49	14	US-10-143-090-187
8	260	100.0	70	10	US-09-866-050A-313
9	260	100.0	71	10	US-09-884-276-116
10	260	100.0	71	10	US-09-984-271-116
11	260	100.0	106	11	US-09-903-190-90
12	260	100.0	106	11	US-09-978-160A-746
13	260	100.0	106	15	US-10-264-237-2068

14	260	100.0	112	9	US-09-925-301-1413	Sequence 1413, Ap
15	260	100.0	132	15	US-10-276-774-2348	Sequence 2348, Ap
16	158	60.8	30	9	US-09-984-245-324	Sequence 324, App
17	158	60.8	30	10	US-09-966-262-324	Sequence 324, App
18	158	60.8	30	10	US-09-983-966-324	Sequence 324, App
19	158	60.8	30	14	US-10-059-395-324	Sequence 324, App
20	158	60.8	30	14	US-10-143-090-324	Sequence 324, App
21	97	37.3	20	9	US-09-984-245-322	Sequence 322, App
22	97	37.3	20	10	US-09-966-262-322	Sequence 322, App
23	97	37.3	20	10	US-09-983-966-322	Sequence 322, App
24	97	37.3	20	14	US-10-059-395-322	Sequence 322, App
25	97	37.3	26	14	US-10-143-090-322	Sequence 322, App
26	90.5	34.8	126	16	US-10-425-114-64499	Sequence 64499, A
27	90.5	34.8	153	16	US-10-767-701-40770	Sequence 40770, A
28	86.5	33.3	113	16	US-10-437-963-158231	Sequence 158231, A
29	86.5	33.3	118	16	US-10-425-114-67570	Sequence 67570, A
30	86.5	33.3	120	15	US-10-425-114-45342	Sequence 45342, A
31	86.5	33.3	132	15	US-10-425-114-67236	Sequence 67236, A
32	74.5	28.7	101	15	US-10-424-599-265641	Sequence 265641, A
33	74.5	28.7	118	15	US-10-425-114-49227	Sequence 49227, A
34	61	23.5	360	15	US-10-104-047-3609	Sequence 3609, Ap
35	61	23.5	406	15	US-10-104-047-3609	Sequence 3609, Ap
36	61	23.5	406	15	US-10-220-475A-18	Sequence 47, App1
37	61	23.5	406	15	US-10-072-012-447	Sequence 447, App
38	61	23.5	416	15	US-10-072-012-124	Sequence 124, App
39	59	22.7	416	15	US-10-072-012-124	Sequence 124, App
40	58	22.3	391	15	US-10-369-493-1081	Sequence 1081, Ap
41	58	22.3	391	16	US-10-104-047-2798	Sequence 2798, Ap
42	56.5	21.7	80	16	US-10-437-963-145446	Sequence 145446, A
43	56.5	21.7	80	16	US-10-437-963-161241	Sequence 161241, A
44	56.5	21.7	539	16	US-10-437-963-154140	Sequence 154140, A
45	55.5	21.3	533	16	US-10-437-963-147568	Sequence 147568, A
			511	14	US-10-080-170-641	Sequence 641, App

ALIGNMENTS

RESULT 1
US-09-984-245-187
Sequence 187, Application US/09984245
Patent No. US20020165374A1
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OR INVENTION: 87 Human Secreted Proteins
FILE REFERENCE: P2004P1
CURRENT FILING DATE: 2001-10-29
CURRENT APPLICATION NUMBER: US/09/984,245
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 09/154,707
PRIOR FILING DATE: 1997-03-21
PRIOR APPLICATION NUMBER: PCT/US98/05311
PRIOR FILING DATE: 1998-03-19
PRIOR APPLICATION NUMBER: US 60/041,277
PRIOR FILING DATE: 1997-03-21
PRIOR APPLICATION NUMBER: US 60/042,344
PRIOR FILING DATE: 1997-03-21
PRIOR APPLICATION NUMBER: US 60/041,276
PRIOR FILING DATE: 1997-03-21
PRIOR APPLICATION NUMBER: US 60/041,281
PRIOR FILING DATE: 1997-03-21
PRIOR APPLICATION NUMBER: US 60/048,094
PRIOR FILING DATE: 1997-05-30
PRIOR APPLICATION NUMBER: US 60/048,350
PRIOR FILING DATE: 1997-05-30
PRIOR APPLICATION NUMBER: US 60/048,188
PRIOR FILING DATE: 1997-05-30
PRIOR APPLICATION NUMBER: US 60/048,135
PRIOR FILING DATE: 1997-05-30
PRIOR APPLICATION NUMBER: US 60/050,937
PRIOR FILING DATE: 1997-05-30
PRIOR APPLICATION NUMBER: US 60/048,187
PRIOR FILING DATE: 1997-05-30
PRIOR APPLICATION NUMBER: US 60/048,099
PRIOR FILING DATE: 1997-05-30

```
/ PRIOR APPLICATION NUMBER: US 60/048,352
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,186
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,069
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,095
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,131
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,096
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,355
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,160
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,351
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,154
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/054,804
/ PRIOR FILING DATE: 1997-08-05
/ PRIOR APPLICATION NUMBER: US 60/056,370
/ PRIOR FILING DATE: 1997-08-19
/ PRIOR APPLICATION NUMBER: US 60/060,862
/ PRIOR FILING DATE: 1997-10-02
/ NUMBER OF SEQ ID NOS: 343
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 187
/ LENGTH: 49
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-984-245-187
```

```
Query Match          100.0%; Score 260; DB 9; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.5e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MNLGMIFSMCGMLKLMKMCAMVAAYCSPISFANSRSSSDTKOMSSFM 49
Db 1 MNLGMIFSMCGMLKLMKMCAMVAAYCSPISFANSRSSSDTKOMSSFM 49
```

```
RESULT 2
US-09-984-276-180
/ Sequence 180, Application US/09984276
/ Publication No. US20030017500A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: 71 Human Secreted Proteins
/ FILE REFERENCE: P2030P1
/ CURRENT APPLICATION NUMBER: US/09/984,276
/ PRIOR FILING DATE: 2001-10-29
/ PRIOR APPLICATION NUMBER: 09/482,273
/ PRIOR FILING DATE: 2000-01-13
/ PRIOR APPLICATION NUMBER: 60/092,921
/ PRIOR FILING DATE: 1998-07-15
/ PRIOR APPLICATION NUMBER: 60/092,922
/ PRIOR FILING DATE: 1998-07-15
/ PRIOR APPLICATION NUMBER: 60/092,956
/ PRIOR FILING DATE: 1998-07-15
/ NUMBER OF SEQ ID NOS: 267
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 180
/ LENGTH: 49
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-984-276-180
```

```
Query Match          100.0%; Score 260; DB 10; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.5e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MNLGMIFSMCGMLKLMKMCAMVAAYCSPISFANSRSSSDTKOMSSFM 49
Db 1 MNLGMIFSMCGMLKLMKMCAMVAAYCSPISFANSRSSSDTKOMSSFM 49
```

```
RESULT 3
US-09-984-271-180
/ Sequence 180, Application US/09984271
/ Publication No. US20030040088A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: 71 Human Secreted Proteins
/ FILE REFERENCE: P2030P1
/ CURRENT APPLICATION NUMBER: US/09/984,271
/ PRIOR FILING DATE: 2001-10-29
/ PRIOR APPLICATION NUMBER: 09/482,273
/ PRIOR FILING DATE: 2000-01-13
/ PRIOR APPLICATION NUMBER: PCT/US99/15849
/ PRIOR FILING DATE: 1999-07-14
/ PRIOR APPLICATION NUMBER: 60/092,921
/ PRIOR FILING DATE: 1998-07-15
/ PRIOR APPLICATION NUMBER: 60/092,922
/ PRIOR FILING DATE: 1998-07-15
/ PRIOR APPLICATION NUMBER: 60/092,956
/ PRIOR FILING DATE: 1998-07-15
/ NUMBER OF SEQ ID NOS: 267
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 180
/ LENGTH: 49
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-984-271-180
```

```
Query Match          100.0%; Score 260; DB 10; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.5e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MNLGMIFSMCGMLKLMKMCAMVAAYCSPISFANSRSSSDTKOMSSFM 49
Db 1 MNLGMIFSMCGMLKLMKMCAMVAAYCSPISFANSRSSSDTKOMSSFM 49
```

```
RESULT 4
US-09-966-262-187
/ Sequence 187, Application US/09966262
/ Publication No. US20030050461A1
/ GENERAL INFORMATION:
/ APPLICANT: Young et al.
/ TITLE OF INVENTION: 87 Human Secreted Proteins
/ FILE REFERENCE: P2004P1
/ CURRENT APPLICATION NUMBER: US/09/966,262
/ PRIOR FILING DATE: 2001-10-01
/ PRIOR APPLICATION NUMBER: US 09/154,707
/ PRIOR FILING DATE: 1998-09-17
/ PRIOR APPLICATION NUMBER: PCT/US98/05311
/ PRIOR FILING DATE: 1998-03-19
/ PRIOR APPLICATION NUMBER: US 60/041,277
/ PRIOR FILING DATE: 1997-03-21
/ PRIOR APPLICATION NUMBER: US 60/042,344
/ PRIOR FILING DATE: 1997-03-21
/ PRIOR APPLICATION NUMBER: US 60/041,276
/ PRIOR FILING DATE: 1997-03-21
/ PRIOR APPLICATION NUMBER: US 60/041,281
/ PRIOR FILING DATE: 1997-03-21
/ PRIOR APPLICATION NUMBER: US 60/048,094
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,350
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,188
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,135
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/050,937
```



```

; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 187
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-966-262-187
Query Match      100.0%; Score 260; DB 10; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.5e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      1 MNLGMIFFSMCGMLTKLKWCAMVAVYCSFISFANSRSSEDTKQMSSEFM 49
Db      1 MNLGMIFFSMCGMLTKLKWCAMVAVYCSFISFANSRSSEDTKQMSSEFM 49

RESULT 5
US-09-983-966-187
; Sequence 187, Application US/09983966
; Publication No. US20030060619A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/09/983,966
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
```

```

; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 187
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-983-966-187
Query Match      100.0%; Score 260; DB 10; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.5e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      1 MNLGMIFFSMCGMLTKLKWCAMVAVYCSFISFANSRSSEDTKQMSSEFM 49
Db      1 MNLGMIFFSMCGMLTKLKWCAMVAVYCSFISFANSRSSEDTKQMSSEFM 49

RESULT 6
US-10-059-395-187
; Sequence 187, Application US/10059395
; Publication No. US20030018180A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/10/059,395
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US/09/966,262
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: US 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
```

```

; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 187
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-059-395-187

```

```

Query Match      100.0%; Score 260; DB 14; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.5e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MNLGMITSMCGMLKMKCAWAVYCSFISFANSRSSSEDTKOMSSFM 49
Db      1 MNLGMITSMCGMLKMKCAWAVYCSFISFANSRSSSEDTKOMSSFM 49

RESULT 7
US-10-143-090-187
; Sequence 187, Application US/10143090
; Publication No. US20030069406a1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: PZ004P1
; CURRENT APPLICATION NUMBER: US/10/143,090
; CURRENT FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: 09/154,707

```

```

; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 187
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-143-090-187

```

```

Query Match      100.0%; Score 260; DB 14; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.5e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MNLGMITSMCGMLKMKCAWAVYCSFISFANSRSSSEDTKOMSSFM 49
Db      1 MNLGMITSMCGMLKMKCAWAVYCSFISFANSRSSSEDTKOMSSFM 49

```

```

RESULT 8
US-09-866-050A-313
; Sequence 313, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Ornuet, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 313
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Human
US-09-866-050A-313

```

```

Query Match      100.0%; Score 260; DB 10; Length 70;
Best Local Similarity 100.0%; Pred. No. 6.5e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MNLGMITSMCGMLKMKCAWAVYCSFISFANSRSSSEDTKOMSSFM 49
Db      1 MNLGMITSMCGMLKMKCAWAVYCSFISFANSRSSSEDTKOMSSFM 49

```

```

RESULT 9
US-09-984-276-116
; Sequence 116, Application US/09984276
; Publication No. US20030017500A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: PZ030P1
; CURRENT APPLICATION NUMBER: US/09/984,276
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482,273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 116
; LENGTH: 71

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (71)
; OTHER INFORMATION: Xaa equals stop translation
US-09-984-276-116
```

```
Query Match          100.0%; Score 260; DB 10; Length 71;
Best Local Similarity 100.0%; Pred. No. 6.6e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 MNLIGMIFSMCGMLKIKKWCAMVAVYCSFISFANSRSSEDTKQMMSSFM 49
Db      1 MNLIGMIFSMCGMLKIKKWCAMVAVYCSFISFANSRSSEDTKQMMSSFM 49
```

```

RESULT 10
US-09-984-271-116
; Sequence 116, Application US/09984271
; Publication No. US20030040088A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/984,271
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482,273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: PCT/US99/15849
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 116
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (71)
; OTHER INFORMATION: Xaa equals stop translation
US-09-984-271-116
```

```
Query Match          100.0%; Score 260; DB 10; Length 71;
Best Local Similarity 100.0%; Pred. No. 6.6e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 MNLIGMIFSMCGMLKIKKWCAMVAVYCSFISFANSRSSEDTKQMMSSFM 49
Db      1 MNLIGMIFSMCGMLKIKKWCAMVAVYCSFISFANSRSSEDTKQMMSSFM 49
```

```

RESULT 11
US-09-903-190-90
; Sequence 90, Application US/09903190
; Publication No. US20030162176A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Mline Edwards, Jean-Baptiste
; APPLICANT: Duclerc, Aymeric
; APPLICANT: Bougueterec, Lydie
; TITLE OF INVENTION: Complementary DNAs
; FILE REFERENCE: GENSET.021A
; CURRENT APPLICATION NUMBER: US/09/903,190
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US/09/247,155A
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/074,121
```

```

; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-04
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: Patent.pm
; SEQ ID NO 90
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -71...-1
US-09-903-190-90
```

```
Query Match          100.0%; Score 260; DB 10; Length 106;
Best Local Similarity 100.0%; Pred. No. 1e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 MNLIGMIFSMCGMLKIKKWCAMVAVYCSFISFANSRSSEDTKQMMSSFM 49
Db      37 MNLIGMIFSMCGMLKIKKWCAMVAVYCSFISFANSRSSEDTKQMMSSFM 85
```

```

RESULT 12
US-09-978-360A-746
; Sequence 746, Application US/09978360A
; Publication No. US20040110939A1
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Mline
; APPLICANT: Duclerc, Aymeric
; APPLICANT: Bougueterec, Lydie
; APPLICANT: Jobert, Severin
; APPLICANT: Clusel, Catherine
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: 56.USA.CTP
; CURRENT APPLICATION NUMBER: US/09/978,360A
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: US 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: US 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 60/099,273
; PRIOR FILING DATE: -09-04
; PRIOR APPLICATION NUMBER: US 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: US 09/215,435
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: PCT/IB98/02122
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: US 09/247,155
; PRIOR FILING DATE: 1999-02-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 810
; SOFTWARE: Patent.pm
; SEQ ID NO 746
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -71...-1
US-09-978-360A-746
```

Query Match 100.0%; Score 260; DB 11; Length 106;
Best Local Similarity 100.0%; Pred. No. 1e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMCGMLKLKWCAMVAAYVCSFISFANSRSSSEDTKOMSSFM 49
DB 37 MNLIGMIFSMCGMLKLKWCAMVAAYVCSFISFANSRSSSEDTKOMSSFM 85

RESULT 13

US-10-264-237-2068
; Sequence 2068, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Brise et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA31PI
; CURRENT APPLICATION NUMBER: US/10/264,237
; PRIOR FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2068
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-237-2068

Query Match 100.0%; Score 260; DB 15; Length 106;
Best Local Similarity 100.0%; Pred. No. 1e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMCGMLKLKWCAMVAAYVCSFISFANSRSSSEDTKOMSSFM 49
DB 37 MNLIGMIFSMCGMLKLKWCAMVAAYVCSFISFANSRSSSEDTKOMSSFM 85

RESULT 14

US-09-925-301-1413
; Sequence 1413, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1413
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1413

Query Match 100.0%; Score 260; DB 9; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.1e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMCGMLKLKWCAMVAAYVCSFISFANSRSSSEDTKOMSSFM 49
DB 43 MNLIGMIFSMCGMLKLKWCAMVAAYVCSFISFANSRSSSEDTKOMSSFM 91

RESULT 15
US-10-276-774-2348

; Sequence 2348, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: No. US20040053245A1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-010
; CURRENT APPLICATION NUMBER: US/10/276,774
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 2348
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-276-774-2348

Query Match 100.0%; Score 260; DB 15; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.3e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMCGMLKLKWCAMVAAYVCSFISFANSRSSSEDTKOMSSFM 49
DB 63 MNLIGMIFSMCGMLKLKWCAMVAAYVCSFISFANSRSSSEDTKOMSSFM 111

Search completed: March 6, 2005, 08:19:20
Job time : 94 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 6, 2005, 07:47:05 ; Search time 53 Seconds
(without alignments)
88.955 Million cell updates/sec

Title: US-10-664-358-549

Perfect score: 260

Sequence: 1 MNVLGMIFSMGGLMKLKKWC.....ISFANSRSBDTKQWSSFM 49

Scoring table:

BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.79.*
2: PIR1.*
3: PIR2.*
4: PIR3.*
5: PIR4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	203	78.1	113	2	T16603
2	73.5	28.3	107	2	T45621
3	59	22.7	432	2	F64484
4	58	22.3	971	2	E96794
5	58	22.7	426	2	E87488
6	57	21.9	103	2	S59846
7	57	21.9	391	2	G90266
8	57	21.9	489	2	T26382
9	56	21.5	275	2	S61179
10	55.5	21.3	511	2	C70803
11	55	21.2	659	2	T01520
12	54	20.8	520	2	T12487
13	54	20.8	821	2	T19705
14	53.5	20.6	217	2	D61277
15	53.5	20.6	495	2	D86442
16	53	20.4	209	2	E82388
17	53	20.4	853	2	S34682
18	52.5	20.2	84	2	S62868
19	52.5	20.2	260	2	E81349
20	52	20.0	214	2	G97008
21	52	20.0	231	2	E90887
22	52	20.0	231	2	D85730
23	52	20.0	231	2	S11429
24	52	20.0	523	2	T07834
25	52	20.0	636	2	A86248
26	52	20.0	758	2	T26595
27	51.5	19.8	758	2	T23276
28	51.5	19.8	520	2	S45753
29	51	19.6	283	2	T21622

30	51	19.6	525	2	E86363	probable thiamin b
31	51	19.6	556	1	OXBY34	DNA endonuclease I
32	51	19.6	854	2	AC2309	hypothetical prote
33	50.5	19.4	154	2	G98112	hypothetical prote
34	50.5	19.4	179	2	S38468	hypothetical prote
35	50.5	19.4	217	2	AG1640	hypothetical prote
36	50.5	19.4	229	2	T03405	probable chitinase
37	50.5	19.4	304	2	T33632	probable cytochrome
38	50.5	19.4	484	2	T34251	hypothetical prote
39	50.5	19.4	572	2	T32636	hypothetical prote
40	50.5	19.4	813	2	UC5785	ATP-dependent RNA
41	50	19.2	174	2	D82629	hypothetical prote
42	50	19.2	325	2	S56788	hypothetical prote
43	50	19.2	370	2	H70423	oxygen-independent
44	50	19.2	405	2	E82626	hypothetical prote
45	50	19.2	668	2	E96777	probable anion exc

ALIGNMENTS

RESULT 1
T16603
hypothetical protein K10B2.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T16603
R:Miller, N.
submitted to the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans coemid K10B2.
A:Reference number: Z18545
A:Accession: T16603
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-113 <MIL>
A:Cross-references: UNIPROT:Q09993; EMBL:U28730; NID:G860694; PID:G860698; PIDN:AAA68261
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CBSP:K10B2.4
A:Introns: 15/2; 53/3; 86/2
C:Superfamily: Caenorhabditis elegans hypothetical protein K10B2.4

Query Match 78.1%; Score 203; DB 2; Length 113;
Best Local Similarity 69.4%; Pred. No. 4.3e-20;
Matches 34; Conservative 12; Mismatches 3; Indels 0; Gaps 0;

QY 1 MNVLGMIFSMGGLMKLKKWCAMVAVVCSFISFANSRSBDTKQWSSFM 49
DB 38 MNVLGMIFSMGGLMKLKKWCAMVAVVCSFISFANSRSBDTKQWSSFM 86

RESULT 2
T45621
hypothetical protein F13G24.160 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C:Accession: T45621
R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.-J.; Voet, M.; Robben, J.; Volckaert, G.; Ba
submitted to the Protein Sequence Database, December 1999
A:Reference number: Z23009
A:Accession: T45621
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-107 <BEV>
A:Cross-references: UNIPROT:Q9SD88; EMBL:AL133421
A:Experimental source: cultivar Columbia; BAC clone F13G24
C:Genetics:
A:Map position: 5
A:Introns: 53/3; 87/2
A:Note: F13G24.160
C:Superfamily: Caenorhabditis elegans hypothetical protein K10B2.4

Query Match 28.3%; Score 73.5; DB 2; Length 107;


```

A:Accession: G90266
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-391 <KUN>
A:Cross-references: UNIPROT.Q97208; GB:AEO06641; NID:g131814326; PIDN:AAK41390.1; GSPDB:G
C:Gene: S801138

Query Match
Beat Local Similarity 23.9%; Score 57; DB 2; Length 391;
Matches 11; Conservative 16; Mismatches 19; Indels 0; Gaps 0;

      4 LGMTFMSMGLMLKLKMCAMVAVYCSFISFANSRSSEDTKOMMSFM 49
      ||::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      229 LAVFLGLTLNVSLGYSPFIISAFVAGVAFANSVKNSKIKEISTLT 274

RESULT 8
T26382
hypothetical protein Y105CSB.f - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T26382
R:McMurray, A.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z20208
A:Accession: T26382
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-489 <WIL>
A:Cross-references: UNIPROT.Q9NAM8; EMBL.NL110479; PIDN:CAB54354.1; CESP:Y105CSB.f
C:Genetic8:
A:Experimental source: clone Y105CSB

A:Gene: CESP:Y105CSB.f
A:Introns: 58/3; 79/3; 114/1; 148/3; 176/2; 233/3; 316/3; 432/1
C:Superfamily: Caenorhabditis elegans hypothetical protein Y105CSB.f

Query Match
Beat Local Similarity 34.6%; Score 57; DB 2; Length 489;
Matches 9; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

      6 MIFSMCGMLKLKMCAMVAVYCSPTS 31
      ::||::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      87 LVFDGCDLVLPIDICIWRFQSFSS 112

RESULT 9
S61179
probable membrane protein YDR384c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein D9481.20
C:Species: Saccharomyces cerevisiae
C:Date: 23-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
C:Accession: S61179; S69668
R:Ding, H.
submitted to the EMBL Data Library, June 1995
A:Description: The sequence of S. cerevisiae cosmid y481.
A:Reference number: S61159
A:Accession: S61179
A:Molecule type: DNA
A:Residues: 1-575 <DIN>
A:Cross-references: UNIPROT.Q12359; EMBL.U28373; NID:g849184; PIDN:AAE4820.1; PID:g84927
A:Experimental source: strain S288C (AB972)
R:Dieckrich, F.S.
submitted to the EMBL Data Library, July 1995
A:Description: The sequence of S. cerevisiae cosmids y481, y509, y926, y461, and lambda da
A:Reference number: S69665
A:Accession: S69668
A:Molecule type: DNA
A:Residues: 1-275 <DIR>
A:Cross-references: EMBL.U32274; NID:g927313; PIDN:AAE4826.1; PID:g927317; MIPS.YDR384c
C:Genetic8:
A:Cross-references: SGD:S0002792

```

A:Map position 4R
C:Superfamily: Saccharomyces cerevisiae probable membrane protein FUN34
C:Keywords: transmembrane protein
F:89-105/Domain: transmembrane #status predicted <TM1>
F:124-140/Domain: transmembrane #status predicted <TM2>
F:182-199/Domain: transmembrane #status predicted <TM3>
F:208-224/Domain: transmembrane #status predicted <TM4>

Query Match 21.5%; Score 56; DB 2; Length 275;
Best Local Similarity 37.9%; Pred. No. 6.3;
Matches 11; Conservative 4; Mismatches 6; Indels 8; Gaps 1;

QY 7 IFSMGIMLKMKCAWVAYCSPISFANS 35
Db 242 ILISC-----CGMYSLICSVSPSNS 262

RESULT 10
C70803
hypothetical protein R3877 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: C70803
R:Colle, S.T.; Brosch, R.; Parthill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felkwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
N:ature 353, 537-544, 1998
A:Author: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: C70803
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-511 <COL>
A:Cross-references: UNIPROT:O69741; GB:AL022120; GB:AL123456; NID:g3261558; PIDN:CAA1796
C:Genetic source: UNIPROT:O69741; GB:AL022120; GB:AL123456; NID:g3261558; PIDN:CAA1796
A:Gene: RV3877

Query Match 21.3%; Score 55.5; DB 2; Length 511;
Best Local Similarity 42.9%; Pred. No. 13;
Matches 9; Conservative 5; Mismatches 4; Indels 3; Gaps 1;

QY 5 GMIFSMGIMLKL--KWCAM 22
Db 401 GLITVCGFRSRLYERWCAM 421

RESULT 11
T01520
hypothetical protein T10M3.18 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 09-Jul-2004
C:Accession: T01520; T01431
R:Johnson, R.F.; de la Bastide, M.; Lodi, M.; Hofman, J.; Haegawa, A.; Gnoj, L.; Gottt.
Martensen, F.; McCombie, W.
A:Submitted to the EMBL Data Library, May 1997
A:Description: The sequence of the Arabidopsis thaliana T10M3 BAC.
A:Reference number: Z14346
A:Accession: T01520
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-659 <JOH>
A:Cross-references: UNIPROT:O04257; EMBL:AF001308; NID:g2104523; PID:g3912924
A:Experimental source: cultivar Columbia
R:Kalicki, J.; Gibson, A.
A:Submitted to the EMBL Data Library, August 1998
A:Description: The sequence of A. thaliana TZH3.
A:Reference number: Z14324
A:Accession: T01431
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 190-659 <KAL>

A:Cross-references: EMBL:AF075597; NID:g3298610; PID:g3377807
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 4S
A>Note: T10M13.18; T2H3.14

Query Match 21.2%; Score 55; DB 2; Length 659;
Best Local Similarity 40.7%; Pred. No. 19;
Matches 11; Conservative 6; Mismatches 8; Indels 2; Gaps 2;

QY 2 NLGMIFFSMGMLTKKWCAMVAAYVC 27
DB 306 NLGGLVSVCGVCRQKVDW-TWGSYSC 331

RESULT 12

T12487
hypothetical protein DKFP56K0247.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 23-Jul-1999

C/Accession: T12487

R:Ottenwälder, B.; Obermaier, B.; Mewes, H.W.; Gaassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, June 1999

A:Reference number: Z17528

A:Accession: T12487

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-520 <OTT>

A:Cross-references: EMBL:AL080226

A:Experimental source: fetal kidney; clone DKFP56K0247

C:Genetics:
A>Note: DKFP56K0247.1

Query Match 20.8%; Score 54; DB 2; Length 520;
Best Local Similarity 34.9%; Pred. No. 21;
Matches 15; Conservative 6; Mismatches 16; Indels 6; Gaps 2;

QY 3 LIGMIFSMGMLTKKWCAMVAAYCSFISFAN-----SRSEPT 41
DB 321 LPEVLNDCGNKTHSKWACWTPV--TNIELCNORASTSGDT 361

RESULT 13

T19705

hypothetical protein C34C6.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C/Accession: T19705

R:Pericy, C.

submitted to the EMBL Data Library, October 1995

A:Reference number: Z19167

A:Accession: T19705

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-821 <MTL>

A:Cross-references: UNIPROT:Q18423; EMBL:Z66494; PIDD:CAA91257.1; GSPDB:GN00020; CESP:CA

A:Experimental source: clone C34C6

C:Genetics:
A:Gene: CESP:C34C6.2

A:Map position: 2

A:Introns: 50/3; 79/2; 139/1; 301/3; 334/1; 387/1; 469/3; 614/1; 658/1; 716/1; 778/3

C:Superfamily: Caenorhabditis elegans hypothetical protein C34C6.2

Query Match 20.8%; Score 54; DB 2; Length 821;
Best Local Similarity 46.2%; Pred. No. 31;
Matches 12; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 17 LKWCAMVAAYCSFISFANSRSSEPTK 42
DB 113 LKCMOVNHNHSLTUSFATSRREYETLK 138

RESULT 14

AG1277
hypothetical protein lmo1623 [imported] - Listeria monocytogenes (strain EGD-e)

C:Species: Listeria monocytogenes

C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C/Accession: AG1277

R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,

.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgey, O.; Entian, K.D.; Feihl, H.

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Matcournam, A.; Mat

ok, C.; Schlueter, T.; Simoes, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland,

A>Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AG1277

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-217 <GLA>

A:Cross-references: UNIPROT:Q6Y6R1; GB:NC_003210; PIDD:CA09701.1; PID:g16411059; GSPDB:G

A:Experimental source: strain EGD-e

C:Genetics:
A:Gene: lmo1623

Query Match 20.6%; Score 53.5; DB 2; Length 217;
Best Local Similarity 31.5%; Pred. No. 11;
Matches 17; Conservative 7; Mismatches 17; Indels 13; Gaps 2;

QY 1 MNLGMIFFSMGMLTKKWCAMVAAYCSFISFANSRSSEPTKQWSSPM 49
DB 11 IGLGLFIFMTGWTGAKMQRPDHYNSVIRVGIT-----DKTIVYISIL 56

RESULT 15

D86442
probable amino acid permease [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C/Accession: D86442

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.P.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Marzalli,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: AB6141; MUID:21016719; PMID:11130712

A:Accession: D86442

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-495 <STO>

A:Cross-references: UNIPROT:Q9C6S5; GB:AB005172; NID:g11136719; PIDD:AA031300.1; GSPDB:G

C:Genetics:
A:Map position: 1

Query Match 20.6%; Score 53.5; DB 2; Length 495;
Best Local Similarity 32.6%; Pred. No. 23;
Matches 15; Conservative 8; Mismatches 10; Indels 13; Gaps 3;

QY 3 LIGMIFSMGMLTKKWCAMVAAY-----YC-----SFISFANSR 36
DB 359 LIGLIFSSAGSVL-LGMLSFQELVAENLLCYGMLLEFIAFVRMR 403

Search completed: March 6, 2005, 08:09:18
Job time: 57 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 6, 2005, 06:42:10 ; Search time 123 seconds
(without alignments)
203.999 Million cell updates/sec

Title: US-10-664-358-549

Perfect score: 260

Sequence: 1 MNLGMIFFMCGMLKLMKC.....ISFANSRSSEDTKQMSSEFM 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	260	100.0	104	2	Q9BV13
2	260	100.0	106	1	CGEO_HUMAN
3	260	100.0	106	2	Q6Q7K0
4	260	100.0	106	2	Q6ZWX0
5	259	99.6	106	2	Q6DKM8
6	259	99.6	106	2	Q6NTW8
7	259	99.6	106	2	Q6PC46
8	221	85.0	130	2	Q7QIP7
9	219	84.2	108	1	U139_MANSE
10	217	83.5	102	2	Q6XIW7
11	217	83.5	108	1	U139_DROME
12	203	78.1	113	1	U139_CABEL
13	90.5	34.8	153	2	Q8LKT8
14	86.5	33.3	93	2	Q9AUI6
15	80	30.8	98	2	Q86HE5
16	73.5	28.3	107	1	U139_ARATH
17	65	25.0	95	1	YA38_SCHPO
18	61	23.5	352	2	Q6ICM2
19	61	23.5	406	1	S144_HUMAN
20	61	23.5	415	2	Q660E5
21	60	23.1	671	2	Q92KX3
22	59.5	22.9	439	2	Q6F6C6
23	59.5	22.9	1802	2	Q9VQD8
24	59	22.7	209	2	Q639A5
25	59	22.7	432	1	YE79_MERVA
26	59	22.7	971	2	Q9C9T5
27	58.5	22.5	322	1	SNS2_HUMAN
28	58	22.3	214	2	Q9N9T4
29	58	22.3	426	2	Q9A6Z6
30	57.5	22.1	387	2	Q83D79
31	57	21.9	298	2	Q6G1Q0

32	57	21.9	358	2	Q8G5K8	Q8G5K8 bifidobacte
33	57	21.9	391	2	Q97Z08	Q97Z08 sulfobact
34	57	21.9	447	2	Q9NAM8	Q9NAM8 caenorhabdi
35	57	21.9	491	2	Q9U622	Q9U622 drosophila
36	57	21.9	572	2	Q811T9	Q811T9 plasmodium
37	56.5	21.7	503	2	Q7ZYD3	Q7ZYD3 xenopus lae
38	56.5	21.7	543	2	Q7UHR4	Q7UHR4 rhodospirill
39	56.5	21.7	583	2	Q67M66	Q67M66 oryza sativ
40	56.5	21.7	846	2	Q9ZOU6	Q9ZOU6 brassica ol
41	56.5	21.7	1520	2	Q6ZPP4	Q6ZPP4 mus musculu
42	56.5	21.7	4903	1	MLL3_MOUSE	Q8BTH4 mus musculu
43	56	21.5	275	1	YD84_YEAST	Q12355 baccharomyc
44	56	21.5	403	1	S144_MOUSE	Q810E9 mus musculu
45	56	21.5	454	2	Q9MJ58	Q9MJ58 loligo blee

ALIGNMENTS

RESULT 1	ID	Q9BV13	PRELIMINARY;	PRT;	104 AA.
AC	Q9BV13	Q9BV13			
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
DT	01-JUN-2003	(TREMBLrel. 24, Last annotation update)			
DE	PRD008 protein (Fragment).				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISUE=Eye;				
RA	Strauberg R.;				
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; BC001192; AA01192.1; -				
DR	InterPro: IPR005351; UPF0139.				
DR	Pfam; PF03669; UPF0139.1.				
DR	ProDom; PD073147; UPF0139; 1.				
FT	NON TER				
SQ	SEQUENCE	104 AA;	11850 MW;	DOB525EA862800A1 CRC64;	
Query Match		100.0%;	Score 260;	DB 2;	Length 104;
Best Local Similarity		100.0%;	Pred. No. 8.4e-27;		
Matches	49;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MNLGMIFFMCGMLKLMCAWVAVYCSFISFANSRSSEDTKQMSSEFM 49			
DB	35	MNLGMIFFMCGMLKLMCAWVAVYCSFISFANSRSSEDTKQMSSEFM 83			
RESULT 2					
CGEO_HUMAN					
ID	CGEO_HUMAN	STANDARD;	PRT;	106 AA.	
AC	Q9Y2E4;				
DT	16-OCT-2001	(Rel. 40, Created)			
DT	16-OCT-2001	(Rel. 40, Last sequence update)			
DT	29-MAR-2004	(Rel. 43, Last annotation update)			
DE	UPF0139 protein CGI-140 (PTD008) (My006 protein).				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	MEDLINE=20272150; PubMed=10810093; DOI=10.1101/gr.10.5.703;				
RT	"Identification of novel human genes evolutionarily conserved in				
RT	Cenorhaditis elegans by comparative proteomics.";				
RL	Genome Res. 10:703-713(2000).				
RN	[2]				
RP	SEQUENCE FROM N.A.				

```

RC TISSUE=Pituitary tumor;
RA Song H., Peng Y., Dai M., Huang Q., Mao Y., Zhang Q., Mao M., Fu G.,
RA Luo M., Chen J., Hu R.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RA Mao Y.M., Xie Y., Zhou Z.X.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the UPF0139 (CGI-140) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF151898; AAD34135.1; -
DR EMBL; AF078861; AAD44493.1; -
DR EMBL; AF059620; AAG43119.1; -
DR InterPro: IPR005351; UPF0139.
DR Pfam: PF03669; UPF0139; 1.
DR ProDom: PD073147; UPF0139; 1.
KW Transmembrane.
FT TRANSMEM 80
SQ SEQUENCE 106 AA; 12068 MW; 7D2BDADF279071B CRC64;

Query Match
Best Local Similarity 100.0%; Score 260; DB 1; Length 106;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMCGIMLKMKCAWVAAYVCSFISFANSRSSSEDTKQWMSFM 49
DB 37 MNLIGMIFSMCGIMLKMKCAWVAAYVCSFISFANSRSSSEDTKQWMSFM 85

RESULT 3
QY 0607K0 PRELIMINARY; PRT; 106 AA.
AC 0607K0;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE CGI-140.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN
RP SEQUENCE FROM N.A.
RA Lee H.Y., Cui X.S., Shin M.L., Jeong Y.J., Hwang K.C., Kim N.H.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY553921; AAS76542.1; -
DR InterPro: IPR005351; UPF0139.
DR Pfam: PF03669; UPF0139; 1.
DR ProDom: PD073147; UPF0139; 1.
DR SEQUENCE 106 AA; 11942 MW; 0CB2D4249AD258E7 CRC64;

Query Match
Best Local Similarity 100.0%; Score 260; DB 2; Length 106;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMCGIMLKMKCAWVAAYVCSFISFANSRSSSEDTKQWMSFM 49
DB 37 MNLIGMIFSMCGIMLKMKCAWVAAYVCSFISFANSRSSSEDTKQWMSFM 85

RESULT 4
QY 06ZMX0 PRELIMINARY; PRT; 106 AA.
AC 06ZMX0;

```

```

DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Mus musculus 13 days embryo liver cDNA. RIKEN full-length enriched
DE library, clone:2510002B07 product:PROTEIN CGI-140 (PROTEIN PT008)
DE homolog (cDNA sequence BC056474).
CN Name=BC056474;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=liver;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=liver;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=liver;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=liver;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugihara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=liver;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=liver;
RA Adachi J., Aizawa K., Akahira S., Akiyama T., Arai A., Aono H.,
RA Arikawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuo M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imocani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Niehi K., Nomura K., Numata R., Ono M.,
RA Okazaki Y., Okido T., Owa C., Saito K., Saito R., Sakai K., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tegawa Y., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.

```

RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Tothiyaki S., Carninci P., Prange C.,
 RA Raha S.S., Loguélino N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalka U., Smalhus D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (JUG-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK010893; BAB27250.1; -
 DR EMBL; BC056474; AAH56474.1; -
 DR InterPro; IPR005351; UPF0139.
 DR Pfam; PF03669; UPF0139.1.
 DR ProDom; PD073147; UPF0139.1.
 SQ SEQUENCE 106 AA; 12068 MW; 7D2BDAF9279071B CRC64;

Query Match 100.0%; Score 260; DB 2; Length 106;
 Best Local Similarity 100.0%; Pred. No. 8.6e-27;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLGMIFFSMCGMLTKKMCANVAVYCSFISFANSRSSEDTKQMSFPM 49
 Db 37 MNLGMIFFSMCGMLTKKMCANVAVYCSFISFANSRSSEDTKQMSFPM 85
 ID Q6DKM8 PRELIMINARY; PRT; 106 AA.
 AC Q6DKM8;
 DT 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE MGC82186 protein.
 GN Name=MGC82186;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Tothiyaki S., Carninci P., Prange C.,
 RA Raha S.S., Loguélino N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalka U., Smalhus D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalka U., Smalhus D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [12]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT Initiative.";
 RL Dev. Dyn. 225:384-391(2002).
 RN [13]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Gerhard D.S.;
 RL Submitted (May-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC071133; AAH71133.1; -
 DR InterPro; IPR005351; UPF0139.
 DR Pfam; PF03669; UPF0139.1.
 DR ProDom; PD073147; UPF0139.1.
 SQ SEQUENCE 106 AA; 11885 MW; 169F889AD4BBB90 CRC64;

Query Match 99.6%; Score 259; DB 2; Length 106;
 Best Local Similarity 98.0%; Pred. No. 1.2e-26;
 Matches 48; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLGMIFFSMCGMLTKKMCANVAVYCSFISFANSRSSEDTKQMSFPM 49
 Db 37 MNLGMIFFSMCGMLTKKMCANVAVYCSFISFANSRSSEDTKQMSFPM 85
 ID Q6NTW8 PRELIMINARY; PRT; 106 AA.
 AC Q6NTW8;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE MGC81480 protein.
 GN Name=MGC81480;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Tothiyaki S., Carninci P., Prange C.,
 RA Raha S.S., Loguélino N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalka U., Smalhus D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RT Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC068836; AAH68836.1; -.
DR InterPro; IPR005351; UPF0139.
DR Pfam; PF03669; UPF0139; 1.
DR Prodom; PD073147; UPF0139; 1.
SQ SEQUENCE 106 AA; 11923 MW; 5CF064CCD2C1EDF CRC64;

Query Match 99.6%; Score 259; DB 2; Length 106;
Best Local Similarity 98.0%; Pred. No. 1.2e-26;
Matches 48; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MNLGMIFSMCGMLKIKWCAMVAVYCSFISFANSRSSEDTKOMSSFM 49
DB 37 MNLGMIFSMCGMLKIKWCAMVAVYCSFISFANSRSSEDTKOMSSFM 85
|||||
RESULT 7
O6PC46 PRELIMINARY; PRT; 106 AA.
AC O6PC46;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein zgc:73111.
GN ORFNames=zgc:73111;
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wild-type; TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Hopkins R.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Ditchenko L., Marins K., Farmer A.A., Rabin G.M., Hong L.,
RA Stepieton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshitsuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhally S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Valladao D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fehey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A., Green E.D., Dickson M.C.,
RA Blakeley R.W., Touchman J.W., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Wild-type; TISSUE=Eye;
RA Strausberg R.,

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC059477; AAH59477.1; -.
DR ZFIN; ZDB-GENE-040426-1674; zgc:73111.
DR InterPro; IPR005351; UPF0139.
DR Pfam; PF03669; UPF0139; 1.
DR Prodom; PD073147; UPF0139; 1.
KW Hypothetical protein.
SQ SEQUENCE 106 AA; 12202 MW; 7DEA7587EDD80A2 CRC64;

Query Match 99.6%; Score 259; DB 2; Length 106;
Best Local Similarity 98.0%; Pred. No. 1.2e-26;
Matches 48; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MNLGMIFSMCGMLKIKWCAMVAVYCSFISFANSRSSEDTKOMSSFM 49
DB 37 MNLGMIFSMCGMLKIKWCAMVAVYCSFISFANSRSSEDTKOMSSFM 85
|||||
RESULT 8
O7QIP7 PRELIMINARY; PRT; 130 AA.
ID O7QIP7
AC O7QIP7;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGCP3421 (Fragment).
GN Name=agCG54409; ORFNames=ENSNANG00000012639;
OS Anopheles gambiae str. PST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB0108807; EAA04290.1; -.
DR InterPro; IPR005351; UPF0139.
DR Pfam; PF03669; UPF0139; 1.
DR Prodom; PD073147; UPF0139; 1.
FT NON TER 1
SQ SEQUENCE 130 AA; 14648 MW; 8F3BF7A8B421F67 CRC64;

Query Match 85.0%; Score 221; DB 2; Length 130;
Best Local Similarity 79.6%; Pred. No. 1.7e-21;
Matches 39; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

OY 1 MNLGMIFSMCGMLKIKWCAMVAVYCSFISFANSRSSEDTKOMSSFM 49
DB 58 MNLGMIFSMCGMLKIKWCAMVAVYCSFISFANSRSSEDTKOMSSFM 106
|||||
RESULT 9
U139 MANSE STANDARD; PRT; 108 AA.
ID U139 MANSE
AC O9U516;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 29-MAR-2004 (Rel. 43, Last annotation update)
DE Hypothetical UPF0139 protein pmemA27.
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Sphingioidea;
OC Sphingidae; Sphinginae; Manduca.
OX NCBI_TaxID=7130;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Antenna;
RX MEDLINE=20099029; PubMed=10620045;
RA Robertson H.M., Matros R., Sears C.R., Todres E.Z., Walden K.K.O.,

RA Nardí J.B.:
 RT "Diversity of odorant binding proteins revealed by an expressed
 RT sequence tag project on male *Manduca sexta* moth antennae.",
 RL Insect Mol. Biol. 8:501-518 (1999).
 CC -1- SIMILARITY: Belongs to the UPF0139 (CGI-140) family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).
 CC -----
 DR EMBL: AF117572; AAI16694.1; -
 DR InterPro: IPR005351; UPF0139.
 DR Pfam: PF03669; UPF0139; 1.
 DR ProDom: PD073147; UPF0139; 1.
 KM Hypothetical protein; Transmembrane.
 FT TRANSMEM 77 Potential.
 SO SEQUENCE 108 AA; 12125 MW; 1C57B0933A798921 CRC64;
 Query Match 84.2%; Score 219; DB 1; Length 108;
 Best Local Similarity 77.6%; Pred. No. 2.6e-21;
 Matches 38; Conservative 8; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MNLGIMFSWCGMLKMKCAWVAVYCSFISFANSRSSEDTKQMSSEFM 49
 DB 34 MNLGIMFSWCGMLMKMKCAWVAVYCSFISFANSRVSDDTKQIVSSEFM 82
 RESULT 10
 Q6XIW7 PRELIMINARY; PRT; 102 AA.
 ID Q6XIW7
 AC Q6XIW7
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Similar to Drosophila melanogaster CG10674 (Fragment).
 OS Drosophila yakuba (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7245;
 OK
 RN
 RP
 RX MEDLINE=22887302; PubMed=14525923; DOI=10.1101/gr.1311003;
 RA Domazet-Loso T., Tautz D.;
 RT "An evolutionary analysis of orphan genes in *Drosophila*.";
 RL Genome Res. 13:2213-2219(2003).
 DR EMBL: AY31712; AAR09735.1; -
 DR InterPro: IPR005351; UPF0139.
 DR Pfam: PF03669; UPF0139; 1.
 DR ProDom: PD073147; UPF0139; 1.
 FT NON_TER 1
 FT NON_TER 102
 SO SEQUENCE 102 AA; 11397 MW; B13E02BD0915D79 CRC64;
 Query Match 83.5%; Score 217; DB 2; Length 102;
 Best Local Similarity 77.6%; Pred. No. 4.6e-21;
 Matches 38; Conservative 8; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MNLGIMFSWCGMLKMKCAWVAVYCSFISFANSRSSEDTKQMSSEFM 49
 DB 37 MNLGIMFSWCGMLMKMKCAWVAVYCSFISFANSRVSDDTKQIVSSEFM 85
 RESULT 11
 U139_DROME STANDARD; PRT; 108 AA.
 ID U139_DROME
 AC Q9VR08;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 25-JAN-2005 (Rel. 46, Last annotation update)
 DE Hypothetical UPF0139 protein CG10674.
 GN ORFNames=CG10674;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 OK
 RN
 RP
 RX STRAIN=Berkelley;
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne U.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoikins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.F., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kechum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Matei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskens D.R., Paclob J.M.,
 RA Palazolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reiner R., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN
 RP
 RX GENOME REANNOTATION.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Miura S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
 RA Suth C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.F.,
 RA Beccencourt B.R., Celnik S.E., de Grey A.D.N., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 CC -1- SIMILARITY: Belongs to the UPF0139 (CGI-140) family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).
 CC -----

DR EMBL; AB003567; AAF50797.1; -
 DR InLac; Q9VRJ8; -
 DR FLYbase; FBgn0035592; CG10674.
 DR InterPro; IPR005351; UPF0139.
 DR Pfam; PF03669; UPF0139; 1.
 DR ProDom; PD073147; UPF0139; 1.
 DR Hypothetical protein; Transmembrane.
 FT TRANSMEM 80
 SQ SEQUENCE 108 AA; 12082 MW; 440857655D5238DC CRC64;

Query Match 83.5%; Score 217; DB 1; Length 108;
 Best Local Similarity 77.6%; Pred. No. 4.8e-21;
 Matches 38; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMCGMLTKMKCAWAVVCSFISFANSSSEDTKQMSSEFM 49
 DB 37 MNLIGMIFSMCGMLTKMKCAWAVVCSFISFANSSSEDTKQMSSEFM 85

RESULT 12
 U139_CABEL STANDARD; PRT; 113 AA.

AC Q09993;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Hypothetical UPF0139 protein K10B2.4 in chromosome II.
 GN ORFNames=K10B2.4;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Polidoridae; Caenorhabditis.
 OX NCBI_Taxid=6239;

RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=9069613; PubMed=9851916;
 RG The C. elegans sequencing consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RL Science 282:2012-2018(1998).
 CC -1- SIMILARITY: Belongs to the UPF0139 (CGI-140) family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcement/>
 CC or send an email to license@ebi.ac.uk).

DR EMBL; U28730; AAA68261.1; -
 DR PIR; T16603; T16603
 DR WormBase; WBGene00019607; K10B2.4.
 DR WormPep; K10B2.4; CER2011.
 DR InterPro; IPR005351; UPF0139.
 DR Pfam; PF03669; UPF0139; 1.
 DR ProDom; PD073147; UPF0139; 1.
 DR Hypothetical protein; Transmembrane.
 FT TRANSMEM 81
 SQ SEQUENCE 113 AA; 12764 MW; 0FEED534A710DD50 CRC64;

Query Match 78.1%; Score 203; DB 1; Length 113;
 Best Local Similarity 69.4%; Pred. No. 3.7e-19;
 Matches 34; Conservative 12; Mismatches 3; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMCGMLTKMKCAWAVVCSFISFANSSSEDTKQMSSEFM 49
 DB 38 MNLIGMIFSMCGMLTKMKCAWAVVCSFISFANSSSEDTKQMSSEFM 86

RESULT 13
 OBLKT8 PRELIMINARY; PRT; 153 AA.

AC Q8LKT8;
 DT 01-OCT-2002 (TRENBLrel. 22, Created)
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Hypothetical protein 170F8.12.
 GN Name=170F8.12;
 OS Sorghum bicolor (Sorghum) (Sorghum vulgare).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACCAD clade; Panicoidae; Andropogoneae; Sorghum.
 OX NCBI_Taxid=4558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22341483; PubMed=12454082;
 RA Ramakrishna W., Dubcovsky J., Park Y.J., Busco C., Emberton J.,
 RA Samadpour P., Bennetzen J.L.;
 RT "Different types and rates of genome evolution detected by comparative
 RT sequence analysis of orthologous segments from four cereal genomes";
 RL Genetics 162:1389-1400(2002).
 DR EMBL; AF503433; AA047587.1; -
 DR InterPro; IPR005351; UPF0139.
 DR Pfam; PF03669; UPF0139; 1.
 DR ProDom; PD073147; UPF0139; 1.
 DR Hypothetical protein.
 SQ SEQUENCE 153 AA; 16498 MW; 0A59856FE6FC4CE2 CRC64;

Query Match 34.8%; Score 90.5; DB 2; Length 153;
 Best Local Similarity 36.3%; Pred. No. 0.00053;
 Matches 18; Conservative 12; Mismatches 16; Indels 1; Gaps 1;

QY 4 LGMIFSMCGMLTKMKCAWAVVCSFISFANSSSEDTKQMSSEFM 49
 DB 45 LAVFGVGVMLRKYKCSWIAITFCAGSLANKNFENDIKOLSMAFM 91

RESULT 14
 O9AUV6 PRELIMINARY; PRT; 93 AA.
 ID O9AUV6;
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Hypothetical protein OSUNBA0040E01.6;
 GN Name=OSUNBA0040E01.6;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_Taxid=39947;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Buell C.R., Yuan O., Ouyang S., Moffat K.S., Hill J.N., Burr P.C.,
 RA Heio J., Zismann V., Pai G., Bowman C.L., Fujii C.Y., Vanaken S.E.,
 RA Bowman C.L., Craven B., Uteirack T.R., Khalak H., Feldblum T.V.,
 RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

RN [2]
 RP SEQUENCE FROM N.A.
 RA Buell R.;
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC079887; AAK16184.1; -
 DR Gramene; O9AUV6; -
 DR InterPro; IPR005351; UPF0139.
 DR Pfam; PF03669; UPF0139; 1.
 DR ProDom; PD073147; UPF0139; 1.
 DR Hypothetical protein.
 SQ SEQUENCE 93 AA; 9961 MW; B92A623A0380C7EC CRC64;

Query Match 33.3%; Score 86.5; DB 2; Length 93;
 Best Local Similarity 36.2%; Pred. No. 0.0011;
 Matches 17; Conservative 12; Mismatches 17; Indels 1; Gaps 1;

QY 4 LGMIFSMCGMLTKMKCAWAVVCSFISFANSSSEDTKQMSSEFM 49

Db 47 LAVVFGVLGVMRLRYKVCWMTAIFCAQSLVNMKNFENDLQKLSWAFM 93

RESULT 15

Q06H65 PRELIMINARY; PRT; 98 AA.
AC Q06H65;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypochemical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gloeckner G., Bichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox B., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC116960; AAC5170.1; -.
DR InterPro; IPR005351; UPF0139.
DR Pfam; PF03669; UPF0139; 1.
DR ProDom; PD073147; UPF0139; 1.
KW Hypochemical protein.
SQ SEQUENCE 98 AA; 11011 MW; 1D44628B4806ABDE CRC64;

Query Match 30.8%; Score 80; DB 2; Length 98;
Best local similarity 32.6%; Pred. No. 0.0087;
Matches 15; Conservative 10; Mismatches 21; Indels 0; Gaps 0;

QY 2 NLGMIFFSMGMLDKLKCAMVAVYCSFISFANRSSSEDTKQWMS 47
DB 31 SLRFSIIFGFLGIMLKYKICLWVSAVCCVAVLSNLKSKDSSVRIILS 76

Search completed: March 6, 2005, 08:08:14
Job time : 128 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2005, 04:07:08 ; Search time 4556 Seconds
(without alignments)
11039.621 Million cell updates/sec

Title: US-10-664-358-36

Perfect score: 1038
Sequence: 1 gnaatcgcgcagcagcttaac.....aaaaaaaaaaaaaacgcga 1038

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ha:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pt:*
10: gb_ro:*
11: gb_sts:*
12: gb_gy:*
13: gb_un:*
14: gb_yt:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1035	99.7	1038	6 BD074351	BD074351 50 human
2	1011.6	97.5	125780	2 AL138904	AL138904 Homo sapi
3	1011.6	97.5	141670	2 AC104170	AC104170 Homo sapi
4	1011.6	97.5	164500	2 AL355968	AL355968 Homo sapi
5	1011.6	97.5	169304	2 AL731898	AL731898 Homo sapi
6	207.8	20.0	41322	9 ACC04076	ACC04076 Homo sapi
7	206.2	19.9	251124	9 HUAE000660	HE000660 Homo sapi
8	205.4	19.8	99108	9 AC008666	AC008666 Homo sapi
9	205.4	19.8	176629	9 AC008641	AC008641 Homo sapi
10	203.8	19.6	205819	2 AC106805	AC106805 Homo sapi
11	203.6	19.6	116368	9 HS112F19	HS112F19 Human DNA
12	203	19.6	159490	9 AC110594	AC110594 Homo sapi
13	202.8	19.5	133902	9 AC108026	AC108026 Homo sapi
14	202.4	19.5	133860	9 AL157709	AL157709 Human DNA
15	202.2	19.5	65335	9 AC112698	AC112698 Homo sapi
16	202.2	19.5	162104	9 AC026325	AC026325 Homo sapi
17	202.2	19.5	176599	2 AC114911	AC114911 Homo sapi
18	202	19.5	181188	2 AC067983	AC067983 Homo sapi
19	202	19.5	194385	9 CNS00001	AL049776 Human chr

20	201.6	19.4	172770	2 AC023133	AC023133 Homo sapi
21	201.6	19.4	173053	2 AC129893	AC129893 Homo sapi
22	201.6	19.4	181517	2 AC126327	AC126327 Homo sapi
23	201.4	19.4	29133	2 AC139398	AC139398 Homo sapi
24	201.4	19.4	50068	2 AC139097	AC139097 Homo sapi
25	201.4	19.4	134059	2 AC016284	AC016284 Homo sapi
26	201.4	19.4	165902	2 AC106763	AC106763 Homo sapi
27	201.4	19.4	176104	9 AC007163	AC007163 Homo sapi
28	201.4	19.4	206302	2 AC133125	AC133125 Homo sapi
29	201.2	19.4	76771	9 AL451051	AL451051 Human DNA
30	200.6	19.3	168839	9 AC022826	AC022826 Homo sapi
31	200.4	19.3	74486	9 AC055764	AC055764 Homo sapi
32	200.4	19.3	96202	9 AL138824	AL138824 Human DNA
33	200.4	19.3	150140	9 AC112216	AC112216 Homo sapi
34	200.2	19.3	160378	9 AC010281	AC010281 Homo sapi
35	200	19.3	68196	6 C0870013	C0870013 Sequence
36	200	19.3	138121	9 AC025431	AC025431 Homo sapi
37	200	19.3	209779	9 AC066616	AC066616 Homo sapi
38	199.8	19.2	121158	9 AC092977	AC092977 Homo sapi
39	199.8	19.2	123748	9 HS742C19	HS742C19 Human DNA
40	199.8	19.2	181636	2 AC025108	AC025108 Homo sapi
41	199.6	19.2	35001	9 AC139143	AC139143 Homo sapi
42	199.6	19.2	152156	9 AC018552	AC018552 Homo sapi
43	199.6	19.2	156503	9 AC011511	AC011511 Homo sapi
44	199.6	19.2	161345	2 AC091968	AC091968 Homo sapi
45	199.6	19.2	169786	9 BS000112	BS000112 Pan trogl

ALIGNMENTS

RESULT 1	BD074351	1038 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD074351	50 human	secreted proteins.		
DEFINITION	BD074351.1	GI:22619954			
ACCESSION	BD074351	JP 2001514024-A/56.			
VERSION	JP 2001514024-A/56.				
KEYWORDS	JP 2001514024-A/56.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 1038)				
TITLE	Moore, P.A., Ruben, S.M., Lafleur, D.W., Shi, Y., Rosen, C.A., Olsen, H.S., Ebner, R. and Brewer, L.A.				
JOURNAL	50 human secreted proteins				
COMMENT	Patent: JP 2001514024-A 56 11-SEP-2001;				
	HUMAN GENOME SCIENCES INC				
	OS Homo sapiens (human)				
	PN JP 2001514024-A/56				
	PD 11-SEP-2001				
	PF 03-SEP-1998 JP 2000508394				
	PR 05-SEP-1997 US 60/057626, 05-SEP-1997 US 60/057663 PR				
	05-SEP-1997 US 60/057669, 12-SEP-1997 US 60/058667 PR				
	12-SEP-1997 US 60/058974, 12-SEP-1997 US 60/058973 PR				
	12-SEP-1997 US 60/058666				
	PI PAUL A MOORE, STEVEN W RUBEN, DAVID W LAFLEUR, YANGGU SHI, CRAIG A				
	PI ROSEN,				
	PI HENRIK S OLSEN, REINHARD EBNER, LAURIE A BREWER PC				
	C12N15/09, A61K38/00, A61K48/00, A61P17/02, A61P19/06, A61P19/00, PC				
	A61P19/02,				
	PC A61P29/00, A61P31/04, A61P31/12, A61P35/00, C07K14/435, C07K16/18,				
	PC C12N1/15,				
	PC C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12Q1/68, G01N33/53, C12N15/ PC				
	00', A61K37/02, C12N5/00				
	CC n equals a,t,g, or c				
	FH Key				
	FT source				
	FT 1. .1038				
	FT /organism='Homo sapiens (human)'. FT location/Qualifiers				
FEATURES	Location/Qualifiers				

FEATURES

source 1..1038

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 99.7%; Score 1035; DB 6; Length 1038;
Best Local Similarity 100.0%; Pred. No. 1.4e-191;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GNAATTCGCGACGAGTTAATGTAATAAATATTTCTAATAATGTAATTTAAATGGAAATTGA 60
DB 1 GNAATTCGCGACGAGTTAATGTAATAAATATTTCTAATAATGTAATTTAAATGGAAATTGA 60
QY 61 GCATCATGAAAAAAGCTCTTAAGTGTGAAAAAATATTGTGTACATTTTGGTCACTA 120
DB 61 GCATCATGAAAAAAGCTCTTAAGTGTGAAAAAATATTGTGTACATTTTGGTCACTA 120
QY 121 ATCTTCAATTAATCTTTAGTAACTAATGTTAAGTTAGACAGAGTGGAGTCTTAATAG 180
DB 121 ATCTTCAATTAATCTTTAGTAACTAATGTTAAGTTAGACAGAGTGGAGTCTTAATAG 180
QY 181 TAAATGGAGCTGACGACATGAAATATACATATCTAATATTTGTGACTATCTTAATAG 240
DB 181 TAAATGGAGCTGACGACATGAAATATACATATCTAATATTTGTGACTATCTTAATAG 240
QY 241 AAAATCAGGAATTTCAAAACCTGTTAGTTTGGGTATAGTCAATTTTAAATAGT 300
DB 241 AAAATCAGGAATTTCAAAACCTGTTAGTTTGGGTATAGTCAATTTTAAATAGT 300
QY 301 GCGGTATATTTATACATGATTTGAGCTTTGTGMAAATATTTCCCTGACCTTTATTTA 360
DB 301 GCGGTATATTTATACATGATTTGAGCTTTGTGMAAATATTTCCCTGACCTTTATTTA 360
QY 361 GATGAGATCTACAGTGTGAGGCAACTTATATATCTGTCACTCCATTAAGTGTCAATGTC 420
DB 361 GATGAGATCTACAGTGTGAGGCAACTTATATATCTGTCACTCCATTAAGTGTCAATGTC 420
QY 421 AGACTCATCCCATGCTAAATTTATATGTTGTAAATAGCGCTTTGTAAATAGTGTG 480
DB 421 AGACTCATCCCATGCTAAATTTATATGTTGTAAATAGCGCTTTGTAAATAGTGTG 480
QY 481 AGGTCAATATATCAAGGCTTCAAGGKATTAATATATAAAACCTTGKTTTATTCCTG 540
DB 481 AGGTCAATATATCAAGGCTTCAAGGKATTAATATATAAAACCTTGKTTTATTCCTG 540
QY 541 TGAATAMCGTTTTTCCATGCAAGTTAAATTTCTCAGCCTTTAATTTTATTAAT 600
DB 541 TGAATAMCGTTTTTCCATGCAAGTTAAATTTCTCAGCCTTTAATTTTATTAAT 600
QY 601 ATATAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
DB 601 ATATAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 661 GCTTTTGTAAATTTGTTAATCTGCAAAATCTTACAGCTTCTCAATTTTGGTGTGAGAT 720
DB 661 GCTTTTGTAAATTTGTTAATCTGCAAAATCTTACAGCTTCTCAATTTTGGTGTGAGAT 720
QY 721 GAAATATCTTAATGTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
DB 721 GAAATATCTTAATGTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
QY 781 CTTTGCGGGGCGAGGCGAGTGTGATCACTTAAGTGTGAGATGATGATGATGATGATGATG 840
DB 781 CTTTGCGGGGCGAGGCGAGTGTGATCACTTAAGTGTGAGATGATGATGATGATGATGATG 840
QY 841 ACATGATGAAACGTTTGTCTTAAATAATCAAAATTTAGACAGGCGTGTGACACACA 900
DB 841 ACATGATGAAACGTTTGTCTTAAATAATCAAAATTTAGACAGGCGTGTGACACACA 900
QY 901 TCTGTAATTCAGGTAATCAGAGGCTTAACAGAGAAATTCCTGAAACCTGGAGGAGCAG 960
DB 901 TCTGTAATTCAGGTAATCAGAGGCTTAACAGAGAAATTCCTGAAACCTGGAGGAGCAG 960
```

```
QY 961 AGGTTCAGTACGACCATTTGACCTCCAGCCTGGGCAACAGATGAGACTTGTCTCAAAA 1020
DB 961 AGGTTCAGTACGACCATTTGACCTCCAGCCTGGGCAACAGATGAGACTTGTCTCAAAA 1020
QY 1021 AAAAAAAAAAACTCGA 1038
DB 1021 AAAAAAAAAAACTCGA 1038
```

RESULT 2

AL138904/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Plumb, B.

Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

Requests: clonerequest@sanger.ac.uk

On Aug 27, 2000 this sequence version replaced gi:9796269.

Genome Center

Center: Sanger Centre

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

Project Information

Center project name: dj932F4

Summary Statistics

Assembly program: XGAP4; version 4.5

Sequencing vector: plasmid; L08752; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Consensus quality: 118782 bases at least Q40

Consensus quality: 121247 bases at least Q30

Consensus quality: 122835 bases at least Q20

Insert size: 125380; sum-of-contigs

Insert size: 139370; 4.2% error; agarose-fp

Quality coverage: 3.66x in Q20 bases; sum-of-contigs Quality

coverage: 3.42x in Q20 bases; agarose-fp

NOTE: This is a 'working draft' sequence. It currently

consists of 15 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 3279: contig of 3279 bp in length

3280 3379: gap of 100 bp

3380 13996: contig of 10617 bp in length

13997 14096: gap of 100 bp

14097 17758: contig of 3662 bp in length

17759 17858: gap of 100 bp

17859 22625: contig of 4767 bp in length

22626 22725: gap of 100 bp

22726 33140: contig of 10415 bp in length

33141 33240: gap of 100 bp

33241 36089: contig of 2849 bp in length

36090 36189: gap of 100 bp

36190 43716: contig of 7527 bp in length

43717 43817: gap of 100 bp

43817 50092: contig of 6276 bp in length

50093 50193: gap of 100 bp

50193 56140: contig of 5948 bp in length

```
* 56141 56240: gap of 100 bp
* 56241 59540: contig of 3300 bp in length
* 59541 59640: gap of 100 bp
* 59641 82808: contig of 23168 bp in length
* 82809 82908: gap of 100 bp
* 82909 100070: contig of 17162 bp in length
* 100071 100170: gap of 100 bp
* 100171 116654: contig of 16484 bp in length
* 116655 116755: gap of 100 bp
* 116755 119341: contig of 2587 bp in length
* 119342 119442: gap of 100 bp
* 119442 126780: contig of 7339 bp in length.
FEATURES
    source
        1. 126780
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="1"
            /map="P32.3-33"
            /clone="RP5-932F4"
            /clone_1fb="RRC1-5"
            1. 3279
                /note="assembly fragment:00163"
                fragment chain:1"
            3380. 13396
                /note="assembly fragment:00030"
                fragment chain:1"
            14057. 17758
                /note="assembly fragment:00986"
                fragment chain:1"
            17659. 22625
                /note="assembly fragment:01111"
                fragment chain:1"
            22726. 33140
                /note="assembly fragment:00681"
                fragment chain:2"
            33241. 36089
                /note="assembly fragment:01055"
                fragment chain:2"
            36190. 43716
                /note="assembly fragment:00252"
                fragment chain:2"
            43817. 50092
                /note="assembly fragment:00161"
                fragment chain:2"
            50193. 56140
                /note="assembly fragment:00987"
                fragment chain:3"
            56241. 59540
                /note="assembly fragment:00088"
                fragment chain:3"
            59641. 82808
                /note="assembly fragment:00044"
                /note="assembly fragment:00070"
            82909. 100070
                /note="assembly fragment:00064"
                /note="assembly fragment:00287"
            100171. 116654
                /note="assembly fragment:00337.0"
            116755. 119341
                /note="assembly fragment:00940"
            119442. 126780
                /note="assembly fragment:00940"
ORIGIN
Query Match 97.5%; Score 1011.6; DB 2; Length 126780;
Blast Local Similarity 99.1%; Pred. No. 3.5e-187;
Matches 1011; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 16 TTATGATTAATAATTTCTATATGAATTTTATGAGCATTCATAGAAAAA 75
DB 64760 TTAAATGATTAATAATTTCTATATGAATTTTATGAGCATTCATAGAAAAA 64701
QY 76 TGCTCTACTGTTGAAAAATTTATTTGTTACATTTTGTCACATATCTTCAATTA 135
|||||
```

```
DB 64700 TGCTCTACTGTTGAAAAATTTATTTGTTACATTTTGTCACATATCTTCAATTA 64641
QY 136 TTATGATTAATAATTTCTATATGAATTTTATGAGCATTCATAGAAAAA 195
DB 64640 TTATGATTAATAATTTCTATATGAATTTTATGAGCATTCATAGAAAAA 64581
QY 196 AGCATGAAAAATTAACATATCTTAATTTTGTGACTATCTTATAGAAAAATCAGAAATTT 255
DB 64580 AGCATGAAAAATTAACATATCTTAATTTTGTGACTATCTTATAGAAAAATCAGAAATTT 64521
QY 256 CAAAACTTGTTAGTTTATAGGGTATAGTCATTTTAAATGTCGGGTATTTATTTATAC 315
DB 64520 CAAAACTTGTTAGTTTATAGGGTATAGTCATTTTAAATGTCGGGTATTTATTTATAC 64461
QY 316 ATGATTTAGGTTTGTGAAAAATTTTCCCGGACTTTTATTTAGATGAGATCAGT 375
DB 64460 ATGATTTAGGTTTGTGAAAAATTTTCCCGGACTTTTATTTAGATGAGATCAGT 64401
QY 376 GTAGCAAACTTATATATATCTGTCACTCCATTTAGTGCATAGTCAGACTCCCATG 435
DB 64400 GTAGCAAACTTATATATATCTGTCACTCCATTTAGTGCATAGTCAGACTCCCATG 64341
QY 436 CTAATAATTATAGTTGTAAAAATACGCTTTGTAAATGTTGTAGGTCATTATACCA 495
DB 64340 CTAATAATTATAGTTGTAAAAATACGCTTTGTAAATGTTGTAGGTCATTATACCA 64281
QY 496 AGCTTCAGAGKATTAACATTAATAAACCTTGCTTTATCTGTGATMCGGTTT 555
DB 64280 AGCTTCAGAGTATTAACATTAATAAACCTTGCTTTATCTGTGATMCGGTTT 64221
QY 556 TCCATGCAAAAGTTAAAAATTTCTTACGCTTTAAATTTTATTAATATATAGATGAT 615
DB 64220 TCCATGCAAAAGTTAAAAATTTCTTACGCTTTAAATTTTATTAATATATAGATGAT 64161
QY 616 GAGTATGACTACAAACAGAAAAATTAACAGATTTGCTTTGCTTAAATG 675
DB 64160 GAGTATGACTACAAACAGAAAAATTAACAGATTTGCTTTGCTTAAATG 64101
QY 676 TTACCTGACAAAATCTTAAGCCAGTTCTTCAATTTGTTTATAGATGATTAAGTTT 735
DB 64100 TTACCTGACAAAATCTTAAGCCAGTTCTTCAATTTGTTTATAGATGATTAAGTTT 64041
QY 736 TAGTCCAGGGGCTGGGCGCATAGCTGATGCTGCTGCTCCAGTCTTTGCGGGCCAG 795
DB 64040 TAGTCCAGGGGCTGGGCGCATAGCTGATGCTGCTGCTCCAGTCTTTGCGGGCCAG 63981
QY 796 GCAGGTGATCACTTAAGGTCAGAGTTTGAGACCAAGCTGCCAATGTTGAACGTT 855
DB 63980 GCAGGTGATCACTTAAGGTCAGAGTTTGAGACCAAGCTGCCAATGTTGAACGTT 63921
QY 856 GTCTCTACTAAAAATTAACAAAATTTAGACAGGCGTGGGCGACACATCTGTATCCAGCT 915
DB 63920 GTCTCTACTAAAAATTAACAAAATTTAGACAGGCGTGGGCGACACATCTGTATCCAGCT 63861
QY 916 ACTCAGAGGCTTAACACAGAAAAATTTCTTGAACCTGGAGAGCAGAGGTTCCAGTGACC 975
DB 63860 ACTCAGAGGCTTAACACAGAAAAATTTCTTGAACCTGGAGAGCAGAGGTTCCAGTGACC 63801
QY 976 ATTGCACTCCAGCTGGGCAACACAGTGAGACTCTTGTCTCAAAAAAATTTAAATCT 1035
DB 63800 ATTGCACTCCAGCTGGGCAACACAGTGAGACTCTTGTCTCAAAAAAATTTAAATCT 63741
RESULT 3
AC104170/c AC104170 141670 bp DNA linear PRI 23-FEB-2002
LOCUS Homo sapiens chromosome 1 clone RP11-253A20, complete sequence.
DEFINITION AC104170 AL549950
VERSION AC104170.2 GI:18873866
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 141670)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimachak, C., Phelps, K.A., Raymond, C. and Haugen, E.D.
Direct Submission
Unpublished
2 (bases 1 to 141670)
Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.
Direct Submission
Submitted (05-DEC-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 141670)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimachak, C., Phelps, K.A., Raymond, C. and Haugen, E.D.
Direct Submission
Submitted (23-FEB-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On Feb 23, 2002 this sequence version replaced gi:17352434.

Center: University of Washington Genome Center
Center Code: UMGC
Web site: <http://www.genome.washington.edu>
Contact: uwgctgs@u.washington.edu
Drafting Center: SC

Project Information
Center project name: chr-1
Center clone name: RP11-253A20 (sc0660)

Summary Statistics
Sequencing vector: plasmid, L08752, 100% of reads
Chemistry: Dye-terminator BT, 50% of reads
Chemistry: Dye-terminator Big Dye, 50% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 141542 bases at least Q40
Consensus quality: 141661 bases at least Q30
Consensus quality: 141670 bases at least Q20
Insert size: 141884; sum-of-contigs
Quality coverage: 8.5x in Q20 bases; sum-of-contigs

Overlapping Sequences:
5': RP5-932P4 (UMGC:sc0224) AL138904
3': Mapping in progress

Sequence Quality Assessment:
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:
This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

ECORI HindIII BglII

SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt
8696	8789	2027	1998	8506	8441
6	<800	6382	6455	2067	2063
199	<800	512	<800	6929	7240
1588	1583	449	<800	145	<800
10718	10649	7395	7424	13457	13095
651	<800	2977	3004	920	933
5486	5384	1037	1022	1035	1032
4358	4277	4491	4481	15120	14801
453	<800	743	<800	3640	3664
5526	5529	3989	3943	6500	6702
8754	8789	14808	14844	1744	1716
3850	3835	8390	8521	952	933
13263	13308	2696	2710	3927	3892
1302	1373	4219	4162	733	750
221	<800	10947	10800	1618	1595
3182	3221	2158	2142	336	<800
8583	8789	1478	1452	2269	22672
2881	2895	3621	3657	4557	4603
2857	2895	1683	1639	1517	1484
3858	3835	1017	1022	750	750
2049	2051	3136	3207	5719	5693
1828	1812	6377	6455	546	<800
728	<800	540	<800	4139	4124
3103	3095	1891	1890	10543	10438
2820	2895	9321	9186	2613	2638
2571	2641	456	<800	7957	8067
183	<800	8747	8822	293	<800
1422	1373	4056	4039	7593	7563
136	<800	2370	2478	5232	5219
4113	4141	5078	5021	4616	4603
2446	2497	22	<800	90	<800
1212	1162	870	880	344	<800
6024	5993	325	<800	3861	3892
36	<800	566	<800		
4066	4013	135	<800		

```
-----
1339 1373 5115 5021
-----
768 <800 9156 9186
-----
9943 9786 4300 4272
-----
6286 6162 2420 2478
-----
317 <800 3736 3657
-----
2147 2184 732 739
-----
1442 1373
-----
3025 3095
-----
871 891
-----
1170 1162
-----
2718 2641
-----
1173 1162
-----
FEATURES
SOURCE Location/Qualifiers
1..141670
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-253A20"
/clone_1db="RP11 human BAC library 11"
45..212
/misc_feature
/note="Single subclone region"
ORIGIN
Query Match 97.5% Score 101.6; DB 9; Length 141670;
Best Local Similarity 99.1% Pred. No. 3.5e-187;
Matches 1011; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
16 TTAATGTAATAATTTCTATTAATGAATTTAATGGAAATTAAGCATCATAAGAAAAA 75
20783 TTAATGTAATAATTTCTATTAATGAATTTAATGGAAATTAAGCATCATAAGAAAAA 20724
76 TGTCTTACTGTTGAAAAATTTATTTGTTACATTTGGTCAACTATCTTCAATTACTT 135
20723 TGTCTTACTGTTGAAAAATTTATTTGTTACATTTGGTCAACTATCTTCAATTACTT 20664
136 TTAGTACTATATATGTTAGTTGTAACGAGTGCAGCTTTATATATGTAATGCGAGTAC 195
20663 TTAGTACTATATATGTTAGTTGTAACGAGTGCAGCTTTATATATGTAATGCGAGTAC 20604
196 AGCATGAATAATACATATCTATATATTTGTAAGTACTATCTTATAGGAAATCAGAAATTT 255
20603 AGCATGAATAATACATATCTATATATTTGTAAGTACTATCTTATAGGAAATCAGAAATTT 20544
256 CAAAACTGTTAGTTTAAAGGTATAGTCAATTTATTAATGTCGGTATATTTATAC 315
20543 CAAAACTGTTAGTTTAAAGGTATAGTCAATTTATTAATGTCGGTATATTTATAC 20484
316 ATGATTTGACGTTGTGAAATATTTTCCCGGACTTTATTTATAGATGAATCTACAGT 375
20483 ATGATTTGACGTTGTGAAATATTTTCCCGGACTTTATTTATAGATGAATCTACAGT 20424
376 GTAGGCAAACTATATATCTGTCAACTCATAGTGCATAGTCAGACTCAGTCCCATG 435
20423 GTAGGCAAACTATATATCTGTCAACTCATAGTGCATAGTCAGACTCAGTCCCATG 20364
436 CTTAAATATATATGTTGTGAAATATGCTTTGTAATATGTTGTGAGTCAATATACCA 495
20363 CTTAAATATATATGTTGTGAAATATGCTTTGTAATATGTTGTGAGTCAATATACCA 20304
```

```
496 AGCTTCAGAGKATTATACATTAATAAAACCTTGSKTTTTATTCTTGTAATACGTTTTT 555
| | | | | : | | | | | : | | | | |
Db 20303 AGCTTCAGAGTATTATACATTAATAAAACCTTGTTTTATTCTTGTAATACGTTTTT 20244
Qy 556 TCCATGCAAAAGTTAAATTTCTTCAGCCTTTAATTTTTTTTATATATATTAAGATGAT 615
| | | | | : | | | | | : | | | | |
Db 20243 TCCATGCAAAAGTTAAATTTCTTCAGCCTTTAATTTTTTTTATATATATTAAGATGAT 20184
Qy 616 GAGTATGACTACAAAACGAGAAAAAATTAACAGATTTGTTTGAGCTTTTGCTAAATG 675
| | | | | : | | | | | : | | | | |
Db 20183 GAGTATGACTACAAAACGAGAAAAAATTAACAGATTTGTTTGAGCTTTTGCTAAATG 20124
Qy 676 TTACCTGCAAAATCTTAGCCAGTTCTTCAATTTTCTTTGATGATGATATCTAGTTT 735
| | | | | : | | | | | : | | | | |
Db 20123 TTACCTGCAAAATCTTAGCCAGTTCTTCAATTTTCTTTGATGATGATGATATCTAGTTT 20064
Qy 736 TAGTCCAGGGGCTGGGCGGATAGCTGATGCTGCTGATCCAGTCTTTGGCGGGCCGAG 795
| | | | | : | | | | | : | | | | |
Db 20063 TAGTCCAGGGGCTGGGCGGATAGCTGATGCTGCTGATCCAGTCTTTGGCGGGCCGAG 20004
Qy 796 GCAGGTGATCATCTTAAGTACAGAGTTTGAGACCAAGCTGCCCAACATGTTGAACGTT 855
| | | | | : | | | | | : | | | | |
Db 20003 GCAGGTGATCATCTTAAGTACAGAGTTTGAGACCAAGCTGCCCAACATGTTGAACGTT 19944
Qy 856 GTCTCTACTAAATAATACAAAATTTAGACAGGCGTGTGAGCACATCTGTATTCAGCT 915
| | | | | : | | | | | : | | | | |
Db 19943 GTCTCTACTAAATAATACAAAATTTAGACAGGCGTGTGAGCACATCTGTATTCAGCT 19884
Qy 916 ACTCAGAGGCTAACACAGAAAAATTCCTTGAACCTGGAGGACAGAGTTGCAGTGAGCC 975
| | | | | : | | | | | : | | | | |
Db 19883 ACTCAGAGGCTAACACAGAAAAATTCCTTGAACCTGGAGGACAGAGTTGCAGTGAGCC 19824
Qy 976 ATTGCATCTCCAGCCTGGGCAACACAGTGAAGCTCTTGTCTCAAAAAAATTTAACT 1035
| | | | | : | | | | | : | | | | |
Db 19823 ATTGCATCTCCAGCCTGGGCAACACAGTGAAGCTCTTGTCTCAAAAAAATTTAACT 19764
```

RESULT 4
ALJ59698 164500 bp DNA linear HTG 10-JUL-2001
LOCUS Homo sapiens chromosome X clone RP11-55316, 3 unordered pieces.
DEFINITION ACCESSION
ALJ59698
VERSION ALJ59698.10 GI:9930965
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Kimberley, A.
Direct Submission
Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
Requests: clonerequests@sanger.ac.uk
On Aug 27, 2000 this sequence version replaced gi:9717116.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: ba55316
----- Summary Statistics
Assembly program: XGAP; version 4.5
Sequencing vector: Plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 163671 bases at least Q40
Consensus quality: 163958 bases at least Q30
Consensus quality: 164134 bases at least Q20
Insert size: 164300; sum-of-contigs
Insert size: 175514; agarose-fp
Quality coverage: 9.44x in Q20 bases; sum-of-contigs Quality
coverage: 8.89x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 23821: contig of 23821 bp in length
* 23822 23921: gap of 100 bp
* 23922 31390: contig of 7469 bp in length
* 31391 31490: gap of 100 bp
* 31491 164500: contig of 133010 bp in length.
* Location/Qualifiers
* 1. 164500
* /organism="Homo sapiens"
* /mol_type="genomic DNA"
* /db_xref="taxon:9606"
* /chromosome="X"
* /clone="RP11-55316"
* /clone_11b="RP11-11.2"
* 1. 23821
* /note="assembly_fragment:02323
* clone_end:SP6
* vector_side:left"
* 23922..31390
* /note="assembly_fragment:01611"
* 31491..164500
* /note="assembly_fragment:01777"

misc_feature
1. 23821
/note="assembly_fragment:02323
clone_end:SP6
vector_side:left"
23922..31390
/note="assembly_fragment:01611"
31491..164500
/note="assembly_fragment:01777"
ORIGIN
Query Match 97.5%; Score 1011.6; DB 2; Length 164500;
Best Local Similarity 99.1%; Pred. No. 3.5e-187;
Matches 1011; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 16 TTAATGTAATAATTTCTCATTAATTAATTTAATGGAATTAAGACATCATAGAAAAA 75
DB 56406 TTAATGTAATAATTTCTCATTAATTAATTTAATGGAATTAAGACATCATAGAAAAA 56465
QY 76 TGTCTTACTGTGAAACATTTTGTACATTTTGGTCACTATCTTCAATAACTT 135
DB 56466 TGTCTTACTGTGAAACATTTTGTACATTTTGGTCACTATCTTCAATAACTT 56525
QY 136 TTAGTAACATTAATGTTAAGTTGACAGTGCAGTCTTAATTAATGAGCACTGAC 195
DB 56526 TTAGTAACATTAATGTTAAGTTGACAGTGCAGTCTTAATTAATGAGCACTGAC 56585
QY 196 AGCATGAAATAACATATCTATATTTTGTGACATCTTATTAAGAAAAACAGAAATT 255
DB 56586 AGCATGAAATAACATATCTATATTTTGTGACATCTTATTAAGAAAAACAGAAATT 56645
QY 256 CAAAACCTTTAGTTTGAAGGTAATGTCACATTTTAATGAGCGGTATTTATAC 315
DB 56646 CAAAACCTTTAGTTTGAAGGTAATGTCACATTTTAATGAGCGGTATTTATAC 56705
QY 316 ATGATTGACGTTGTGMAAATATTTTCCGACCTTTATTTAATGAGATCTAAGT 375
DB 56706 ATGATTGACGTTGTGMAAATATTTTCCGACCTTTATTTAATGAGATCTAAGT 56765
QY 376 GTAGGCAAACTTAATTAATCTGTCACTCATTAAGTCAATGACAGCACTCCCATG 435
DB 56766 GTAGGCAAACTTAATTAATCTGTCACTCATTAAGTCAATGACAGCACTCCCATG 56825
QY 436 CTAATAATATAGTGTAAAAATAGCTTTGTAAATGTTGTAGTCAATATACCA 495
DB 56826 CTAATAATATAGTGTAAAAATAGCTTTGTAAATGTTGTAGTCAATATACCA 56885
QY 496 AGTCTTCAAGGKATTAATTAATAAACCCTTGKTTTATCTTGGAATACCGTTTTT 555
DB 56886 AGTCTTCAAGGKATTAATTAATAAACCCTTGKTTTATCTTGGAATACCGTTTTT 56945
QY 556 TCCATGCAAAAGTTAAATTTCTCAGCCTTAATTTTATTAATTAATTAAGATGTGAT 615

DB 56946 TCCATGCAAAAGTTAAATTTCTCAGCCTTTAATTTTATTAATTAATTAAGATGTGAT 57005
QY 616 GAGTATGACTTACAAAACGAAAAAATTAACAGATTTGCTTGTGCTTAAATG 675
DB 57006 GAGTATGACTTACAAAACGAAAAAATTAACAGATTTGCTTGTGCTTAAATG 57065
QY 676 TTACCTGCAAAAATCTTGCCAGTTCTCATTTTCGTTTGAATGAATGATCTAAGTT 735
DB 57066 TTACCTGCAAAAATCTTGCCAGTTCTCATTTTCGTTTGAATGAATGATCTAAGTT 57125
QY 736 TAGTCAGAGGCTGAGCGCGATAGCTGATGCTGTGCTCCAGTCTTGGGCGGAG 795
DB 57126 TAGTCAGAGGCTGAGCGCGATAGCTGATGCTGTGATTCACAGTCTTGGGCGGAG 57185
QY 796 GCAGGTGATCATCTTAAGTATAGAGTTTGAACACACCTGCGCCCAATGCTGAACGTT 855
DB 57186 GCAGGTGATCATCTTAAGTATAGAGTTTGAACACACCTGCGCCCAATGCTGAACGTT 57245
QY 856 GTCTCTCTAATAAATAACAAAATTAGACAGGCGGTGTGACACATCTGTAATTCAGCT 915
DB 57246 GTCTCTCTAATAAATAACAAAATTAGACAGGCGGTGTGACACATCTGTAATTCAGCT 57305
QY 916 ACTCAGAGAGCTTAACAGAAAAATCTTGAACCTGAGAGGAGAGGTTGACGTAGCC 975
DB 57306 ACTCAGAGAGCTTAACAGAAAAATCTTGAACCTGAGAGGAGAGGTTGACGTAGCC 57365
QY 976 ATTGCACTCCAGCCTGGGCAACACAGTGAAGACTTGTCTCAAAAAAATTAATTAAT 1035
DB 57366 ATTGCACTCCAGCCTGGGCAACACAGTGAAGACTTGTCTCAAAAAAATTAATTAAT 57425

RESULT 5
AL731898/c 169304 bp DNA linear HTG 12-MAY-2002
LOCUS AL731898 Homo sapiens chromosome 1 clone RP11-449B11.
DEFINITION AL731898 GI:20795687
VERSION AL731898.3
KEYWORDS HTG; HTGS_PHASE2; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 169304)
REFERENCE
AUTHORS Harrison E.
TITLE Direct Submision
JOURNAL Submitted (08-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On May 14, 2002 this sequence version replaced gi:20429020.
COMMENT
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: ba449B11
----- Summary Statistics
Assembly program: XGAP4, version 4.5
Sequencing vector: M13; M77815; 54% of reads
Sequencing vector: plasmid; 108752; 45% of reads
Chemistry: Dye-terminator Big Dye; 99% of reads
Chemistry: Dye-terminator; 0% of reads
Consensus quality: 169292 bases at least Q40
Consensus quality: 169299 bases at least Q30
Consensus quality: 169301 bases at least Q20
Insert size: 169304; sum-of-contigs
Insert size: 174813; 2.3% error; agarose-fp
Quality coverage: 18.08x in Q20 bases; sum-of-contigs Quality
coverage: 17.90x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

1 169304; contig of 169304 bp in length.

Location/Qualifiers

source
 1. 169304
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="RP11-449h11"
 /clone_11b="RPCT-11.2"
 1. 169304
 /note="assembly_Fragment:02115
 clone end:T7
 vector_side:right"

ORIGIN

Query Match 97.5%; Score 1011.6; DB 2; Length 169304;
 Best Local Similarity 99.1%; Pred. No. 3.5e-187;
 Matches 1011; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 16 TTAATGTTAAATTTCTTAATGAATTTTAAAGGAAATTAAGCATCTAGAAAAA 75
 Db 139430 TTAATGTTAAATTTCTTAATGAATTTTAAAGGAAATTAAGCATCTAGAAAAA 139371
 QY 76 TGCCTTAACCTGTGAAAAATTTATTTGTTACATTTTGGTCAATATCTTCAATTAATT 135
 Db 139370 TGCCTTAACCTGTGAAAAATTTATTTGTTACATTTTGGTCAATATCTTCAATTAATT 139311
 QY 136 TTAGTAATCTAATAGTTAAGTTGACAGGCGAGCTTATATAGTAATAGCAGCTGAC 195
 Db 139310 TTAGTAATCTAATAGTTAAGTTGACAGGCGAGCTTATATAGTAATAGCAGCTGAC 139251
 QY 196 AGCATGAAAAATTAACATATCTAATATTTTGGCATCTCTATTAAGAAAAATCAGAAATTT 255
 Db 139250 AGCATGAAAAATTAACATATCTAATATTTTGGCATCTCTATTAAGAAAAATCAGAAATTT 139191
 QY 256 CAAACCTGTGTTAGTTTAAAGGTATAGTCATTTTAAATGCGGTATATTTTATAC 315
 Db 139190 CAAACCTGTGTTAGTTTAAAGGTATAGTCATTTTAAATGCGGTATATTTTATAC 139131
 QY 316 ATGATTGACGTTTGTGMAAATATTTTCCCTGCACTTTTATTTAGATGAGATCTACAGT 375
 Db 139130 ATGATTGACGTTTGTGMAAATATTTTCCCTGCACTTTTATTTAGATGAGATCTACAGT 139071
 QY 376 GTAGGCAAACTTATATATCTGTCACTGCTTAGTGTGATGACATCATCCCATG 435
 Db 139070 GTAGGCAAACTTATATATCTGTCACTGCTTAGTGTGATGACATCATCCCATG 139011
 QY 436 CTAATAATATATAGTTGTAATAATAGCTTTGTAATAGTTGTAAGTCAATATACCA 495
 Db 139010 CTAATAATATATAGTTGTAATAATAGCTTTGTAATAGTTGTAAGTCAATATACCA 138951
 QY 496 AGCTTCAGGKATTAATTAATAAACTTGKTTTATCTTGTAATACCGTTTT 555
 Db 138950 AGCTTCAGGKATTAATTAATAAACTTGKTTTATCTTGTAATACCGTTTT 138891
 QY 556 TCCATGCAAAATTAATTTCTTCAAGCTTTAATTTTATTAATATATAGATGATGAT 615
 Db 138890 TCCATGCAAAATTAATTTCTTCAAGCTTTAATTTTATTAATATATAGATGATGAT 138831
 QY 616 GAGTATGACTACAAACAGAAAAATPAACAGATTTGCTTGGCTTTGCTAAATG 675
 Db 138830 GAGTATGACTACAAACAGAAAAATPAACAGATTTGCTTGGCTTTGCTAAATG 138771
 QY 676 TTACTGTGCAAAATCTTAGGACAGTTCTTCAATTTGTTTTGAGATGAAGATTAAGTTT 735
 Db 138770 TTACTGTGCAAAATCTTAGGACAGTTCTTCAATTTGTTTTGAGATGAAGATTAAGTTT 138711

QY 736 TAGTCCAGGGGCTGGGGCCGATAGCTGATGCTGTGATCCAGATCTTTGGGGCCGAG 795
 Db 138710 TAGTCCAGGGGCTGGGGCCGATAGCTGATGCTGTGATCCAGATCTTTGGGGCCGAG 138651
 QY 796 GCAGGTGATCACTTAAGGTTCAGAGTTTGAACACAGCTGCCCCAATGCTGAACGTT 855
 Db 138650 GCAGGTGATCACTTAAGGTTCAGAGTTTGAACACAGCTGCCCCAATGCTGAACGTT 138591
 QY 856 GTCTCTACTAAAAATTAACAAAATTGACAGGCGGTGGGCAACATCTGTATTCAGCT 915
 Db 138590 GTCTCTACTAAAAATTAACAAAATTGACAGGCGGTGGGCAACATCTGTATTCAGCT 138531
 QY 916 ACTCAGAGGCTTAACACAGGAAATTTCTTGAACCTGSGAGGACAGGTTGACAGTACC 975
 Db 138530 ACTCAGAGGCTTAACACAGGAAATTTCTTGAACCTGSGAGGACAGGTTGACAGTACC 138471
 QY 976 ATTCGACTCCAGCCTGGGCAACACAGTGAAGCTCTTGTCTCAAAAAAAAAAACT 1035
 Db 138470 ATTCGACTCCAGCCTGGGCAACACAGTGAAGCTCTTGTCTCAAAAAAAAAAACT 138411

RESULT 6
 AC004076 41322 bp DNA linear PRI 29-JAN-1998
 LOCUS Homo sapiens chromosome 19, cosmid R30217, complete sequence.
 DEFINITION
 AC004076
 VERSION AC004076.1 GI:2822142
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 41322)
 Lamerdin,J.E., McCreedy,P.M., Showronski,E., Adamson,A.W.,
 Burkhardt-Schultz,K., Gordon,L., Kyle,A., Ramirez,M., Stillwgen,S.,
 Garne,J., Dangnan,L., Poundstone,P., Christensen,M.,
 Georgescu,A., Avila,J., Liu,S., Bruce,R., Quan,G., Montgomey,M.,
 Ow,D., Nolan,M., Trong,S., Kobayashi,A., Olsen,A.O. and
 Carrano,A.V.
 Sequence analysis of a 500 kb ZNF gene family- containing human
 contig in 19q13.4
 TITILE Unpublished
 JOURNAL 2 (bases 1 to 41322)
 REFERENCE Lamerdin,J.E.
 AUTHORS Direct Submission
 TITLE Submitted (29-JAN-1998) Joint Genome Institute, Lawrence Livermore
 JOURNAL National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
 COMMENT Map and sequence oriented from centromere to q telomere.

Cosmid R30217 overlaps cosmid R28253 to the left and F18750 to the
 right.

FEATURES
 source
 1. 41322
 /location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="19"
 /map="19q13.4 between D19S303 and ZNF134"
 /clone="R30217"
 /cell_line="SH12-B"
 /clone_11b="LL19NC032 R chromosome 19-specific cosmid"
 /note="cosmid library constructed at LNL from flow-sorted
 chromosomes from hybrid SH12-B, which carries chromosome
 19 as its only human chromosome."
 1. 1431
 /note="BLASTN similarity to AC003002 (80114..81544);
 match: 1, score: 1.8e-294; database searched: nt; Human
 DNA from overlapping chromosome 19-specific cosmid R29515
 and R28253, genomic sequence, complete sequence [Homo
 sapiens]"
 complement(10..98)
 /rpt_family="WStc"
 repeat_region

```
repeat_region      383..680
                    /rpt_family="AluSc"
misc_feature        /note="DPS similarity to AA229025 nc50c11.s1 NCI CGAP_Pr3
                    Homo sapiens cDNA clone IMAGE:1011572. Score: 550
                    identity: 290/300 (96%)."
```

```
repeat_region      1281..1318
                    /rpt_family="POLY_A"
repeat_region      complement(1443..1755)
                    /rpt_family="AluSp"
repeat_region      1917..2204
                    /rpt_family="AluSc"
repeat_region      complement(2735..2777)
                    /rpt_family="(CA)n"
repeat_region      3566..3652
                    /rpt_family="FRAM/FAM"
repeat_region      3671..3719
                    /rpt_family="MER4D"
repeat_region      complement(3720..4019)
                    /rpt_family="AluSp"
repeat_region      4047..4324
                    /rpt_family="MER4D"
misc_feature        4960..5217
                    /note="DPS similarity to multiple ESTs:
                    (4960..5217) AA446441 zw60d10.r1 Soares total fetus NB2HP8
                    9w Homo sapiens cDNA clone 774451 5' similar to contains
                    LTR3.t2 LTR3 repetitive element ;
                    1..259) ; 98% identity.~(4962..5216) AA195132 zrz4b08.r1
                    Soares NBHMPu S1 Homo sapiens cDNA clone 665271 5' similar
                    to contains LTR3.b1 LTR3 repetitive element; (1..253) ; 99%
                    identity.~(4962..5216) W03459 za06e09.r1 Soares melanocyte
                    2NBHM Homo sapiens cDNA clone 291784 5' similar to
                    contains LTR3.b1 LTR3 repetitive element;(1..253) ; 96%
                    identity.
                    and others."
repeat_region      complement(5599..5875)
                    /rpt_family="LINE2"
repeat_region      complement(6952..6980)
                    /rpt_family="AT_rich"
repeat_region      7067..7364
                    /rpt_family="AluSg"
repeat_region      complement(7488..7593)
                    /rpt_family="MIR"
repeat_region      7718..8134
                    /rpt_family="LTR3"
repeat_region      complement(8649..9102)
                    /rpt_family="LINE2"
repeat_region      9114..9386
                    /rpt_family="AluSc"
misc_feature        9573..9676
                    /note="predicted exon, program: Grai12exons_human_1.3,
                    frame: 1, quality: excellent, score: 90.000"
repeat_region      10469..10758
                    /rpt_family="AluSg"
repeat_region      11041..11377
                    /rpt_family="MER74"
misc_feature        11448..11576
                    /note="DPS similarity to (U66561) kruppel-related zinc
                    finger protein [Homo sapiens] (1..43) ; 48% identity.-Other
                    overlapping matches:
                    (11508..11634) predicted exon, program:
                    Grai12exons_human_1.3, frame: 2, quality: excellent,
                    score: 83.000"
misc_feature        12108..12167
                    /note="DPS similarity to (U66561) kruppel-related zinc
                    finger protein [Homo sapiens] (45..63) ; 26% identity."
```

```
repeat_region      12861..12900
                    /rpt_family="MER4D"
misc_feature        12973..13537
                    /note="DPS similarity to (U66561) kruppel-related zinc
                    finger protein [Homo sapiens] (64..249).-Other overlapping
                    matches:
                    (12973..15108) predicted exon, program:
```

```
CDS
                    Grai12exons_human_1.3, frame: 0, quality: excellent,
                    score: 87.000"
                    13033..15108
                    /note="hypothetical ZNF-like protein"
                    /codon_start=1
                    /product="R30217.1"
                    /protein_id="AA87932.1"
                    /db_xref="GI:2822143"
                    /translation="MCSIIILKILILHLAEHDGTHPEOGLYTCAAEHLHQRQIREKLT
                    RSDWRPSPFNHSAHVGEKFTCTOGGKDFPTAASDILQQCVLNSGKLYADTDGEAF
                    QGEONDEPNSSQGGKDFCHQGLFEHOKTHNGERPEFSECEGLFRYNSLITKIQONHA
                    GERPEEGTEYKTFIRKSNLVQHOKIHSEGLSKRSPDIEHOLISRPPECTOGCK
                    AFLTOAHLVGHOKTHNGEOPYECNCKCKPEPMVNSKILRHOKYVITGERRRSCSGFLF
                    MDSPFTLRHRYVHTGERPEFSCICGFRFSRSTLNMHQRVHAKRULYKSCGKAFSL
                    KANVQHLKTHGERPEFTECEKAPVRKSHLVQHOKIHTDASKSDLIQHRIDIR
                    PRPYTSCGKAFLTQAHLVGHOKIHTGERPYCTOCAKAFVKKSHLVQHEKIHIDAF
                    SKRSDLIQHRIDIRPRPYTSCGKAFLTQAHLDHOKIOTGERPECECKFLD
                    SYKLVIHQRIHTGEKPYKSCGKFPYRCTLSDHOKIHTGERPEYKSCGKFPYR
                    KLIIHQVHTGEKPYKSCGKFPYRSTIKKHKICTGEPHEKCEKELPTKESL
                    LIHQSHTEGSPKLABCGNDPDKCNTGQRQKTHTEGSESSKVRKNSLIKH
                    QIIHTGRRP"
                    13592..14095
                    /note="DPS similarity to (U66561) kruppel-related zinc
                    finger protein [Homo sapiens] (250..417)."
```

```
14147..14314
                    /note="DPS similarity to (U66561) kruppel-related zinc
                    finger protein [Homo sapiens] (418..473)."
```

```
14366..15116
                    /note="DPS similarity to (U66561) kruppel-related zinc
                    finger protein [Homo sapiens] (474..726)."
```

```
15272..15565
                    /rpt_family="AluSc"
repeat_region      complement(15566..15720)
                    /rpt_family="L1MB3"
repeat_region      15721..16037
                    /rpt_family="MER1B"
repeat_region      complement(16055..16255)
                    /rpt_family="L1MB3"
repeat_region      16256..17064
                    /rpt_family="MER7B"
repeat_region      complement(17065..17363)
                    /rpt_family="AluSg"
repeat_region      complement(17364..17658)
                    /rpt_family="AluSg"
repeat_region      17659..17849
                    /rpt_family="MER7B"
repeat_region      complement(17853..18015)
                    /rpt_family="L1MB3"
repeat_region      18024..18050
                    /rpt_family="(TAA)n"
repeat_region      complement(18051..18348)
                    /rpt_family="AluY"
repeat_region      complement(18358..19222)
                    /rpt_family="L1MB3"
misc_feature        19980..20439
                    /note="DPS similarity to N20479 yx39g09.s1 Homo sapiens
                    cDNA clone 264160 3'. Score: 900 Identity: 455/460
                    (98%)."
```

```
20474..20521
                    /rpt_family="POLY_A"
repeat_region      complement(20832..20969)
                    /rpt_family="MSTC"
repeat_region      complement(21068..21397)
                    /rpt_family="AluSg"
repeat_region      complement(21402..21590)
                    /rpt_family="MSTC"
repeat_region      21846..22198
                    /rpt_family="MSTC"
repeat_region      complement(22532..22641)
                    /rpt_family="L1MA9"
repeat_region      complement(22738..22856)
                    /rpt_family="(GGA)n"
repeat_region      complement(22871..23168)
```


	/rpt_family="MER58A"
gene	7244..7983
	/gene="TCRAV18S1"
	/pseudo
promoter	7244..7263
	/gene="TCRAV18S1"
	/strand_name="TCRAV18S1"
mRNA	join(<7482..7527,7656..>7944)
	/gene="TCRAV18S1"
	/product="TCRAV18S1"
	/pseudo
V_segment	join(7482..7527,7656..>7944)
	/gene="TCRAV18S1"
	/strand_name="TCRAV18S1"
misc_recomb	/pseudo
	7945..7951
	/gene="TCRAV18S1"
	/strand_name="TCRAV18S1"
	/note="heptamer"
misc_recomb	7952..7974
	/gene="TCRAV18S1"
	/strand_name="TCRAV18S1"
	/note="spacer"
misc_recomb	7975..7983
	/gene="TCRAV18S1"
	/strand_name="TCRAV18S1"
	/note="nonamer"
	complement(8236..8389)
repeat_region	/rpt_family="MER69A"
	complement(8556..8740)
repeat_region	/rpt_family="L1MB7"
	complement(9281..10170)
repeat_region	/rpt_family="L1PA2"
	complement(11002..11230)
repeat_region	/rpt_family="L1"
	11234..11405
	/rpt_family="L1"
gene	11818..12491
	/gene="TCRAV19S1"
	11818..11837
promoter	/gene="TCRAV19S1"
	/strand_name="TCRAV19S1"
	/note="Old name TCRAV12S1"
mRNA	join(<11900..11948,12152..>12452)
	/gene="TCRAV19S1"
	/product="TCRAV19S1"
CDS	/note="Old name TCRAV12S1"
	join(11900..11948,12152..>12452)
	/gene="TCRAV19S1"
	/strand_name="TCRAV19S1"
	/note="Old name TCRAV12S1"
	/codon_start=1
	/product="TCRAV19S1"
	/protein_id="BAB69019.1"
	/db_xref="GI:2358044"
	/translation="MTASTLRVAIVASICVSSMAYERDITTVYLFMYKPPSGELVFLLRRNSFDQNM OVVDASVRYCALSEAA"
V_segment	join(11900..11948,12152..>12452)
	/gene="TCRAV19S1"
	/strand_name="TCRAV19S1"
	/note="Old name TCRAV12S1"
misc_recomb	12453..12459
	/gene="TCRAV19S1"
	/strand_name="TCRAV19S1"
	/note="heptamer"
misc_recomb	12460..12482
	/gene="TCRAV19S1"
	/strand_name="TCRAV19S1"
	/note="spacer"
misc_recomb	12483..12491
	/gene="TCRAV19S1"

	repeat_region	/standard_name="TCRAV19S1"	
		/note="nonamer"	
	repeat_region	13763..13833	
		/rpt_family="MIR2"	
	repeat_region	13879..13946	
		/rpt_family="MIR"	
	repeat_region	complement(19382..19677)	
		/rpt_family="Alusg"	
	repeat_region	21487..21622	
		/rpt_family="MIR"	
	repeat_region	23034..23706	
		/rpt_family="L1ME3A"	
	repeat_region	23955..24354	
		/rpt_family="MSTA"	
Query Match	19.9%; Score 206.2; DB 9; Length 251124;		
Best Local Similarity	81.8%; Pred. No. 1.6e-30;		
Matches 238; Conservative	0; Mismatches 53; Indels 0; Gaps 0;		
Oy	743 GGGGCTGGCGCGGATNAGCTGAATGCCCTGTGTGCCCAATGCTTTGGCGGGCCGAGCAGGNG	802	
Db	114854 GAGGGTGGGCACATTGGCTCATGCTGTATATCCACACTTTGGAGGCTGAGGCGGCG	1149133	
Oy	803 GATCACTTAAGCTCAGAGATTGAGACACAGCCCTGCCCAATGATAAAAGTTGCTCTA	862	
Db	114914 AATCACCTGAGGTCAAGAGCTTGAGACTRAGCCTCGGCCAACATRGTAACAACCCTGCTTTA	11497373	
Oy	863 CTAAAAATACAAAATTTAGACAGCGGTGTGGCACACATCTGTAAATTCAGCTACTCAGG	922	
Db	114974 CTAATAATACAAAATTTAGCCAGGCGATGTGTATGACACCTGTAAATCACAGCTACTCGGG	1150333	
Oy	923 AGCTTAACACAGAGAAATTCCTTTAACTCTGGAGGACAGGTTGACAGTGAACCTATTGCAC	962	
Db	115034 AGCTGAGGCGAGAGATCTGTTGAACCTGGAGGAGGTTGCTGTGAGCCACTGCAC	1150933	
Oy	983 TCACACCTGGGCAACACAGTAGACTCTGTCTCAAAAAAAAAAAAAA	1033	
Db	115094 TTCACCTGGGTAACATGATGAGACTGTCTTCAGAGAAAAA	115144	
RESULT 8			
LOCUS	AC008666	99108 bp	DNA linear PRI 28-FEB-2001
DEFINITION	Homo sapiens chromosome 5 clone CTB-35A8, complete sequence.		
ACCESSION	AC008666		
VERSION	AC008666.5 GI:13162497		
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 99108) DOE Joint Genome Institute and Stanford Human Genome Center. Direct Submission Unpublished 2 (bases 1 to 99108) DOE Joint Genome Institute. Direct Submission Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA 3 (bases 1 to 99108) DOE Joint Genome Institute and Stanford Human Genome Center. Direct Submission Submitted (26-FEB-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA On Feb 28, 2001 this sequence version replaced gi:7709249. Draft Sequence Produced by DOE Joint Genome Institute www.jgi.doe.gov Finishing Completed at Stanford Human Genome Center www.sbgc.stanford.edu Quality: Phrap Quality >=40 99.6% of Sequence; Estimated Total Number of Errors is 0.3. STS Content:		

SHGC-10972 G11217
WI-5180 G02907.
Location/Qualifiers
1. .99108
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTB-35A8"

ORIGIN

Query Match 19.8%; Score 205.4; DB 9; Length 99108;
Best Local Similarity 79.5%; Pred. No. 2.5e-30;
Matches 260; Conservative 0; Mismatches 56; Indels 11; Gaps 1;

Qy 718 GATGAAGTACTTACTGTTTAACTCCAGGGGCTGGCGCGATAGCTGATGCTGTGTGCCA 777
Db 51492 GAGCAGAGAAACAAGTTTCTCCCTGGCTAGGTGTGTGCTCATGCTGTATATCCCA 51551
Qy 778 GTGCTTGGCGGGCGGAGGAGGTGATCACTTAAGTCAGAGATTGAGACCAAGCCTGC 837
Db 51552 GCACCTTGGGAGACTGAGGCGAGGTGATCACTTGAGTCAAGAGTTTGTATACCAAGCTTG 51611
Qy 838 CCAACATGTGTAACGTTGTCTCTACTTAATAAATAAATAATTAGACAGCGGTGTGGCAC 897
Db 51612 CCAACATGTGTAACCCCGTCTGTACTGAAATACAAATAATTAGCAGGCGATGTGTGCAC 51671
Qy 898 ACATCTGTAATTCAGCTACTCAGAGGCTTAACACAGGAAATTCCTTTGAACCTGGGAGG 957
Db 51672 ACACCTGTAGTTCCAGCTACTCAGAGGCTGAGGACAGAGAATCTTTGAACCTGGGAGG 51731
Qy 958 CAGAGGTTGCAGTGA-----GCCATTGCACTCCAGCCTGGGCAACAGAGTGGA 1006
Db 51732 TGGAGGTTCAGTGAAGCTGAGATCATGTCTACTGCAAGTCAGCCTGGGCAACAGAGTGGA 51791
Qy 1007 CTCTTGTCTCAAAAAAAAAAAAAA 1033
Db 51792 CTCTGTCTTAAAAAAAAAAAAAA 51818

RESULT 9
AC008641/c 176629 bp DNA linear PRI 23-AUG-2001
LOCUS Homo sapiens chromosome 5 clone CTB-17P3, complete sequence.
DEFINITION AC008641
ACCESSION AC008641.7 GI:15281183
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 176629)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 176629)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 176629)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (21-OCT-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE 4 (bases 1 to 176629)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Aug 23, 2001 this sequence version replaced gi:10944452.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov

Finishing Completed at Stanford Human Genome Center
www.sbgc.stanford.edu
Quality: Phrap Quality >=40 99.5% of Sequence;
Estimated Total Number of Errors is 0.8.
STS Content:
WI-5180 G02907
SHGC-10972 G11217
WI-12039 G22831.
Location/Qualifiers
1. .176629
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTB-17P3"

ORIGIN

Query Match 19.8%; Score 205.4; DB 9; Length 176629;
Best Local Similarity 79.5%; Pred. No. 2.4e-30;
Matches 260; Conservative 0; Mismatches 56; Indels 11; Gaps 1;

Qy 718 GATGAAGTACTTACTGTTTAACTCCAGGGGCTGGCGCGATAGCTGATGCTGTGTGCCA 777
Db 47412 GAGCAGAGAAACAAGTTTCTCCCTGGCTAGGTGTGTGCTCATGCTGTATATCCCA 47353
Qy 778 GTGCTTGGCGGGCGGAGGAGGTGATCACTTAAGTCAGAGATTGAGACCAAGCCTGC 837
Db 47352 GCACCTTGGGAGACTGAGGCGAGGTGATCACTTGAGTCAAGAGTTTGTATACCAAGCTTG 47293
Qy 838 CCAACATGTGTAACGTTGTCTCTACTTAATAAATAAATAATTAGACAGCGGTGTGGCAC 897
Db 47292 CCAACATGTGTAACCCCGTCTGTACTGAAATACAAATAATTAGCAGGCGATGTGTGCAC 47233
Qy 898 ACATCTGTAATTCAGCTACTCAGAGGCTTAACACAGGAAATTCCTTTGAACCTGGGAGG 957
Db 47232 ACACCTGTAGTTCCAGCTACTCAGAGGCTGAGGACAGAGAATCTTTGAACCTGGGAGG 47173
Qy 958 CAGAGGTTGCAGTGA-----GCCATTGCACTCCAGCCTGGGCAACAGAGTGGA 1006
Db 47172 TGGAGGTTCAGTGAAGCTGAGATCATGTCTACTGCAAGTCAGCCTGGGCAACAGAGTGGA 47113
Qy 1007 CTCTTGTCTCAAAAAAAAAAAAAA 1033
Db 47112 CTCTGTCTTAAAAAAAAAAAAAA 47086

RESULT 10
AC106805/c 205819 bp DNA linear HTG 25-JAN-2002
LOCUS Homo sapiens chromosome 5 clone RP11-507E2, WORKING DRAFT SEQUENCE.
DEFINITION AC106805
ACCESSION AC106805.1 GI:18139355
VERSION HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 205819)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Sequencing of Human Chromosome 5
REFERENCE 2 (bases 1 to 205819)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (12-JAN-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT -----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information

Center Project Name: 601151
Center clone name: RPCI-11_507E2

Summary Statistics

Consensus quality: 202257 bases at least Q40
Consensus quality: 204013 bases at least Q30
Consensus quality: 204448 bases at least Q20
Estimated insert size: 210100; agarose-fp estimation
Estimated insert size: 205419; sum-of-contigs estimation
Quality coverage: 8.87 in Q20 bases; agarose-fp estimation
Quality coverage: 9.07 in Q20 bases; sum-of-contigs estimation
NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 13625: contig of 13625 bp in length
* 13626 13725: gap of unknown length
* 13726 13789: contig of 17564 bp in length
* 13789 31389: gap of unknown length
* 31390 61952: contig of 30563 bp in length
* 61953 62052: gap of unknown length
* 62053 106239: contig of 44247 bp in length
* 106300 106399: gap of unknown length
* 106400 205819: contig of 99420 bp in length.

FEATURES

source
1. 205819
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-507E2"
/clone_1ib="RPCI human BMC library 11"

ORIGIN

Query Match 19.6%; Score 203.8; DB 2; Length 205819;
Best Local Similarity 79.2%; Pred. No. 4.8e-30;
Matches 259; Conservative 0; Mismatches 57; Indels 11; Gaps 1;

718 GATGAAGTACTTACTTATTGTCAGAGGCGCTGCGGCGCATAGCTGATCCCTGATCCCA 777
Db 126190 GAGGAGAGAAACAGTTCCTCGGCTGAGTGTGCTATCCTGTATATCCCA 126131
778 GTGCTTTGGGGGCGGAGGCGAGTGTATCACTTAAGTCAGAGATTGAGACCAAGCTGC 837
Db 126130 GCACCTTTGGGAGACTGAGGCGAGTGTATCACTTAAGTCAGAGATTGATACAGCTCG 126071
838 CCAACATGATGAAAGCTTGTCTCTACTTAAATTAACAAATTAACAGAGCGCTGTGCAC 897
Db 126070 CCAACATGATGAAAGCTTGTCTCTACTTAAATTAACAAATTAACAGAGCGCTGTGCAC 126011
898 ACATCTGTAATTCACGACTACTCAGAGAGGCTTAACAGAAATTCCTTGAACCTGGAGG 957
Db 126010 AACACCTGTAATTCACGACTACTCAGAGAGGCTTAACAGAAATTCCTTGAACCTGGAGG 125951
958 CAGAGCTTGCAGTGA-----GCCATTGCACTTCACCTGGGCAACACAGTGA 1006
Db 125950 TGGAGGTTCAGTGAAGTGAAGTCACTGCTACTGCAAGTCCAGGCAACAGTGA 125891
1007 CTCTGTCTCAAAAAAAAAAAAAA 1033
Db 125890 CTCTGTCTTAATAAAAAAAAAAAAAA 125864

RESULT 11
HS112F19 116368 bp DNA linear PRI 11-MAR-2001
LOCUS
DEFINITION Human DNA sequence from clone RPS-1112F19 on chromosome 20q13.13-13.2 contains the gene for a novel protein similar to SAL1 (Drosophila)-like 1 (LOC57167), ESTs, STSs and a

ACCESSION

AL034420
AL034420.16 GI:5578919
KEYWORDS HTG; CPG island; LOC57167; SAL1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 116368)

REFERENCE

Babbage, A.
Submitted (07-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

AUTHORS

requesters: clonerequests@sanger.ac.uk
On Jul 23, 1999 this sequence version replaced gi:5566333.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

COMMENT

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:
SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 20, constructed by the Sanger Centre Chromosome 20
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr20
RPS-1112F19 is from the library RPCI-5 constructed by the group of
Pleier de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2
This sequence is the entire insert of clone RPS-1112F19.

FEATURES

source

1. 116368
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="RZPD:RPCI704F191112"
/db_xref="taxon:9606"
/chromosome="20"
/map="q13.13-13.2"
/clone="RPS-1112F19"
/clone_1ib="RPCI-5"
1. 174
/note="RIGGRI repeat: matches 1973. 2156 of consensus"
175. 463
/note="AluY repeat: matches 1. 289 of consensus"
471. 614
/note="RIGGRI repeat: matches 1666. 1829 of consensus"
615. 745
/note="Aluub repeat: matches 2. 137 of consensus"
746. 996
/note="RIGGRI repeat: matches 1431. 1692 of consensus"
997. 1541
/note="MER41A repeat: matches 5. 554 of consensus"
1542. 1669
/note="RIGGRI repeat: matches 1299. 1431 of consensus"
1670. 1964
/note="AluSg repeat: matches 1. 294 of consensus"
1965. 2090
/note="RIGGRI repeat: matches 1169. 1299 of consensus"
2091. 2365
/note="AluX repeat: matches 3. 292 of consensus"
2366. 2713

repeat_region /note="TIGER1 repeat: matches 876, .1169 of consensus"
2714, .3017
/note="AluSp repeat: matches 1, .304 of consensus"
3018, .3046
repeat_region /note="TIGER1 repeat: matches 847, .876 of consensus"
3048, .3268
repeat_region /note="AluSp repeat: matches 1, .221 of consensus"
3135, .3226
/note="Single clone region. Single cloned region in Alu"
3290, .3584
repeat_region /note="AluSx repeat: matches 1, .294 of consensus"
3585, .3688
repeat_region /note="MER7A repeat: matches 14, .122 of consensus"
3952, .4186
repeat_region /note="AluY repeat: matches 1, .265 of consensus"
4331, .4591
repeat_region /note="AluSg repeat: matches 1, .260 of consensus"
4592, .4678
repeat_region /note="AluUb repeat: matches 211, .297 of consensus"
4679, .4971
repeat_region /note="AluSp repeat: matches 1, .298 of consensus"
5334, .5357
/note="12 copies 2 mer aa 100% conserved"
6525, .6811
/note="AluSg repeat: matches 2, .289 of consensus"
complement(6825, .25258)
/gene="dJ1112F19.1"
complement(join(6825, .7467,11644, .11924,13805, .15135, 25062, .25258))
/gene="dJ1112F19.1"
/product="dJ1112F19.1 (novel protein similar to SALL1 (sal (Drosophila)-like 1) (LOC57167))"
/note="match: cDNA: Em:146583 Em:Y19265 Em:X97581 Em:U7376 Em:U37251 Em:U37263 Em:U46186 Em:AB007872 Em:AK001666
match: ESTs: Em:A1632569 Em:A1638036 Em:AA417660"
/evidence=not experimental
complement(6825)
/gene="dJ1112F19.1"
complement(6844, .6849)
/gene="dJ1112F19.1"
complement(join(7048, .7467,11644, .11924,13805, .15135, 25062, .25191))
/gene="dJ1112F19.1"
/note="match: protein: Tr:Q91929 Tr:Q98881 Tr:Q62255 Tr:O43340 Tr:O12958 Sw:P51508 Tr:O96395 Tr:Q92117 Sw:P39770 Sw:P39806 Tr:P91639 Tr:Q24219 Tr:O95878 Sw:P18746 Tr:Q90658"
/codon start=1
/evidence=not experimental
/product="dJ1112F19.1 (novel protein similar to SALL1 (sal (Drosophila)-like 1) (LOC57167))"
/protein_id="CAB61485.1"
/db_xref="GI:6456821"
/db_xref="GOA:O9UJ04"
/db_xref="UniProt/Swiss-Prot:O9UJ04"
/translation="MSRRKQAKFOHINSBEOEQOQOQTPEDADAAPAAAGELGAPVNHGVNDEVADEATVKKLRLETHVCEKCCAFESISELHKKCNKPLVEMNDSEGVPSDEFGAVLSHQPTPGSGKCHRENGSGSEDMKEKPDDESVYTERALPTPDQISYLAKKVANTVYTLQALRGTKAVNORADALPAVPANASIPWLEQILCQQOQLOLTEQRIQVNMMAHSHAGSGAGTILKTLGSGAOQVAAVALSOKAGSQSLDALKQAKLPHANIPSNSSSPGLAPTLTKEDGRVLPNWSRPSALLPQAPSVLROSPSTVALDTSKKGKGPPIISADVAKDEBALYHKCYCKSVGTTDSLIQLHRSHTERPFCVCGHRTTKNLIKVHHRHQVIANQLEFQDKVAGNIPYALSVDPDIDBPSLSLDSKPVLYVTISVGLPNLSGTPKDLTGSLGDLQPGPSEBGGPTLPGVPNPNYPAGSGGSGTPEPSEETLKLQVLENDIKATIDN ECLICHRVLSGQSLVLMKQHTRMHGGQIPPTLPBPACDPTGSEBPTUGENSGTAL OHGPCIQCKRTNAVMLQAHTRMGGQIPPTLPBPACDPTGSEBPTUGENSGTAL CHDVISIDIVEIVSOGAPSSSKVPTPLPSIHSASPTLQFAMASIDAFKGPAPFNLDQSGRENGSVSDGLTNDSSSLGWDQYQSRSPDILETTSFQALSPNLSAESI KSKPDGSKAESSENSRTBMBGRSLPSTPIRAPVYKKEVGTGVTGSPNLSGMT PLTAAOPRROAKHGCTRCGNFSSASLQIHERTHGEKFPVNICGRATTTGNLKVHVTWGANNNNSARGRKLAIENTMALLGTGKRVSEIFPEILAPSVNDPVVMQY

repeat_region TSMINGIAVKTNEISVIQSGVPTLPVLSGATSVNNAATVSKMDQSGISADVEKP
SATDGVKQHPHLEENKIAVS"
7578, .7803
/note="MER47A repeat: matches 137, .365 of consensus"
7804, .8106
repeat_region /note="AluSx repeat: matches 1, .305 of consensus"
8107, .8249
repeat_region /note="MER47A repeat: matches 1, .137 of consensus"
8420, .8724
repeat_region /note="AluSc repeat: matches 1, .301 of consensus"
8725, .8972
repeat_region /note="AluJo repeat: matches 136, .290 of consensus"
8873, .9176
repeat_region /note="AluY repeat: matches 1, .296 of consensus"
9177, .9314
repeat_region /note="AluJo repeat: matches 1, .136 of consensus"
9335, .9586
repeat_region /note="L1MD2 repeat: matches 6084, .6341 of consensus"
9596, .9902
repeat_region /note="AluSg repeat: matches 1, .309 of consensus"
9924, .10198
repeat_region /note="AluSx repeat: matches 1, .274 of consensus"
10466, .10589
repeat_region /note="FLAM C repeat: matches 1, .124 of consensus"
10623, .10905
repeat_region /note="AluJo repeat: matches 3, .282 of consensus"
10948, .11191
repeat_region /note="AluJo repeat: matches 81, .289 of consensus"
11192, .11505
repeat_region /note="AluY repeat: matches 1, .306 of consensus"
11506, .11546
repeat_region /note="AluJo repeat: matches 37, .81 of consensus"
complement(14508, .15012)
/gene="dJ1112F19.1"
/note="match: GSS: Em:AQ341439"
complement(14576, .15010)
/gene="match: GSS: Em:AQ341889"
/note="match: GSS: Em:AQ341889"
15190, .15500
repeat_region /note="AluSx repeat: matches 1, .311 of consensus"
15825, .16275
/note="match: STS: Em:N72278"
16000, .16260
repeat_region /note="AluUb repeat: matches 1, .299 of consensus"
16284, .16511
repeat_region /note="AluSx repeat: matches 1, .256 of consensus"
17079, .17363
repeat_region /note="AluSp repeat: matches 1, .312 of consensus"
17440, .17615
/note="MIR repeat: matches 45, .244 of consensus"

Query Match 19.6%; Score 203.6; DB 9; Length 116368;
Best Local Similarity 82.0%; Pred. No. 5,56-30;
Matches 251; Conservative 0; Mismatches 44; Indels 11; Gaps 1;
Qy 739 TCACAGGGCGGCGGCGATGCTATGCTGCTGCCAGTGCCTTGCGGGCGGAGGCA 798
Db 69820 TCAAGGCGCTGGGGGTGGTGTCTGTCTCTTAATCCAGACTTTGGACGCTGAGGTG 69879
Qy 799 GGTGATCACTTAAGGTCAAGAGTTTGAGACCAAGCTGCCCAATGGTGAAGTTGTC 858
Db 69880 GAGGATCACTAGAGTCAAGAGTTTGAGACCAAGCTGCCCAATGGTGAAGCCCTGTC 69939
Qy 859 TCTACTAAAAATACAAAAATTGACAGAGCGGTGTGGACACATCTGTAATTCAGACTACT 918
Db 69940 TCTACTAAAAAACAAAAAATTAGCCGGGGGGGTGGGACACATCTGTAATCCAGACTACT 69999
Qy 919 CAGAGGCTAACAGAAAAAATTCCTTGAACCTGGGAGGACAGAGTTGCACTGA----- 972
Db 70000 CAGAGGCTGACTAGAGAAATCACTTAACCTGGAGAGTGAAGTTGCACTGAGCCGAG 70059
Qy 973 -----GCCATTGCACTCAGGCTGGGCAACACACTGAGACTCTTGTCTCAAAAAAAAAA 1027


```
repeat_region      16215. .16446  
/rpt_family="ALubB"  
repeat_region      complement(16770. .17108)  
/rpt_family="ALubB"  
repeat_region      18252. .18570  
/rpt_family="ALuDo"  
repeat_region      complement(18770. .18989)  
/rpt_family="WIR"  
unSURE  
repeat_region      complement(18897. .19024)  
/note="single clone coverage"  
unSURE  
/rpt_family="LIR28"  
complement(19236. .19253)  
/note "<30 qual SNGL region"  
repeat_region      complement(19244. .20030)  
/rpt_family="LIR1B"  
unSURE  
19818. .20633  
/note="PCR product sequence only"  
repeat_region      complement(20031. .20368)  
/rpt_family="HURS-P2"  
repeat_region      complement(20468. .21174)  
/rpt_family="HURS-P2"  
repeat_region      complement(21174. .24069)  
/rpt_family="HURS-P2"  
24070. .24364  
/rpt_family="ALuSg"  
repeat_region      complement(24365. .25300)  
/rpt_family="HURS-P2"  
repeat_region      complement(25436. .25569)  
/rpt_family="LIR28"  
repeat_region      complement(25594. .26466)  
/rpt_family="LIR1B"  
repeat_region      complement(26936. .27052)  
/rpt_family="LIME1"  
repeat_region      complement(27053. .27348)  
/rpt_family="ALuSc"  
repeat_region      complement(27349. .28223)  
/rpt_family="LIME1"  
repeat_region      complement(28259. .28436)  
/rpt_family="ALuDo"  
repeat_region      complement(28442. .28555)  
/rpt_family="LIME1"  
28556. .28643  
/rpt_family="FLAM_A"  
repeat_region      complement(28644. .28949)  
/rpt_family="ALuSx"  
repeat_region      28950. .28994  
/rpt_family="FLAM_A"  
repeat_region      complement(29006. .29119)  
/rpt_family="LIME1"  
repeat_region      complement(29120. .29418)  
/rpt_family="ALuY"  
repeat_region      complement(29419. .29571)  
/rpt_family="LIME1"  
repeat_region      complement(29583. .29744)  
/rpt_family="LIME1"  
repeat_region      complement(29794. .29997)  
/rpt_family="MER20"  
repeat_region      complement(30024. .30125)  
/rpt_family="LIME1"  
repeat_region      complement(30126. .30263)  
/rpt_family="FLAM_A"  
repeat_region      complement(30264. .30551)
```

	Query Match	19.6%	Score 203	DB 9	Length 159450
	Best Local Similarity	81.1%	Pred. No. 7e-30		
	Matches	236	Conservative	0	Mismatches 55
				Indels	0
				Gaps	0
QY	743	GGGGCTGGGGCCGATGCTGATGCTCTGTGCTGCCAGTCTTTGCGGGGCCGAGCAGAGTG	802		
DB	86873	GGGGCCAGCGCTGTGGCTCAACCTGTAACTTCTTGACACTTTTGGAGGGCCAGGTGGGAG	86881		
QY	803	GATCACTTAAAGTCAGAGATTGAGACCGAGCCTGCCCAACATGTGTGAACGTTGTCTCTA	862		

Db	86613	GATCACCCTGGGCTCAGGAGTTCCAGACCACTCGTCTCAACATGCTAAGGCCCGCTTCTTA	8675
Qy	863	CTAAAAATACAAAATTTAGACAGCGCTGGTGACACATCTGAAATTCACGTA	922
Db	86753	CTAAAAATACAAAATTTAGCAGCGCTGATGACACCCGCTGTAACTCCAGTACTTTAG	8669
Qy	923	AGGCTAACACAGGAAATTCCTTGAACCTTGAGGACGAGAGTTGCAGTGGACATTCGCAC	982
Db	86693	AGGCTAGGACGAGGAATCGGTTGAACCGGGGAACGAGAGTTGGAGTGACCATTCGCAC	8663
Qy	983	TCCAGCCTGGGCAACACAGTGAAGCTCTTGTCTCAAAAAAAAAAAAAAAAAA	1033
Db	86633	TCCAGCCTGGGCAACAGAGTGAAGCTCTTCAAAAAAAAAAAAAAAAAA	86583

RESULT_13	AC108026/c	AC108026	123902 bp	DNA	linear	PRI 13-MAR-2002
LOCUS		Homo sapiens	BAC clone	Rpl1-1E15	from 4,	complete sequence.
DEFINITION		AC108026				
ACCESSION		AC108026.5	GI:19387784			
VERSION						
KEYWORDS		HTG.				
SOURCE		Homo sapiens	(human)			
ORGANISM		Homo sapiens				

REFERENCE	1 (bases 1 to 123902)
AUTHORS	Sulston,J.E. and Waterston,R.
TITLE	Toward a complete human genome sequence
JOURNAL	Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE	99063792

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 123502)
Desai, A. and Cotton, M.
The sequence of Homo sapiens BAC clone RP11-41E15
Unpublished (2001)
3 (bases 1 to 123502)
Waterston, R. H.
Direct Submission
Submitted (24-JUN-2002) Genome Sequencing Center, Washington

REFERENCE
AUTHORS
TITLE
JOURNAL

NC (bases 1 to 123902)
4 (bases 1 to 123902)
Waterston, R. H.
Direct Submission
Submitted (23-FEB-2002) Genome Sequencing Center, Washington

REFERENCE	5 (bases 1 to 123902)
AUTHORS	Waterston,R.
TITLE	Direct Submission
JOURNAL	Submitted (13-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT	On Mar 13, 2002 this sequence version replaced g1.18873868.

Center Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: saplens@watsont.wustl.edu
Summary Statistics
Center project name: H_NH0041E15

NOTE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data subinsets.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPCL11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.edu>
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-811L19; the clone sequenced to the right is RP11-270L13, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-41B15; actual end is at base position 52485 of RP11-270L13.

Polymorphisms exists between AC108026, AC112724 and AC108043. Data from AC112724 was used to finish AC108026.

FEATURES

source

1..123902
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-41B15"
/clone_id="RPCL11"
1..3439
/rpt_family="L1"
3444..3913
/rpt_family="ERV1"
4552..4773
/rpt_family="MER1_type"
5079..5100
/rpt_family="AT_rich"
5248..5393
/rpt_family="L2"
5433..5493
/rpt_family="L2"
6878..6898
/rpt_family="AT_rich"
6918..7076
/rpt_family="L1"
7077..7376
/rpt_family="MER2_type"
7377..7421
/rpt_family="L1"
7424..7956
/rpt_family="L1"
7950..9527
/rpt_family="L1"
9535..9659
/rpt_family="MALR"
9699..9901
/rpt_family="MER1_type"
9902..10204
/rpt_family="MALR"
10216..10241
/rpt_family="(TG)n"
10246..10584
/rpt_family="L1"
10706..11007
/rpt_family="L1"

repeat_region 10989..11236
/rpt_family="L1"
repeat_region 11313..12046
/rpt_family="L1"
repeat_region 12141..12800
/rpt_family="L1"
repeat_region 13250..13342
/rpt_family="MIR"
repeat_region 13361..13391
/rpt_family="L1"
repeat_region 13392..14775
/rpt_family="L1"
repeat_region 14776..14922
/rpt_family="L1"
repeat_region 15274..15366
/rpt_family="L1"
repeat_region 15504..15670
/rpt_family="L1"
repeat_region 16039..16275
/rpt_family="L2"
repeat_region 16651..16785
/rpt_family="L1"
repeat_region 16855..17461
/rpt_family="L1"
repeat_region 17533..18613
/rpt_family="L1"
repeat_region 18738..19039
/rpt_family="L2"
repeat_region 19188..19659
/rpt_family="MALR"
repeat_region 19727..19815
/rpt_family="ERV1"
repeat_region 19890..20160
/rpt_family="L1"
repeat_region 20161..20461
/rpt_family="ALU"
repeat_region 20462..21332
/rpt_family="L1"
repeat_region 21337..21658
/rpt_family="L1"
repeat_region 21667..21728
/rpt_family="L1"
repeat_region 21729..22024
/rpt_family="ALU"
repeat_region 22025..22684
/rpt_family="L1"
repeat_region 22704..22887
/rpt_family="L1"
repeat_region 22888..23174
/rpt_family="ALU"
repeat_region 23175..24161
/rpt_family="L1"
misc_feature 24825..24828
/note="similar to Homo sapiens EST AA173493 (NID:91753661)"
zp02904.t1
25016..25339
/rpt_family="L2"
25340..25520
/rpt_family="MIR"
25784..25812
/rpt_family="(TCTCCC)n"
26401..26501
/rpt_family="L2"
26546..26629
/rpt_family="GA-rich"
27069..27173
/rpt_family="MIR"
28701..28784
/rpt_family="AT_rich"
29258..29502
/rpt_family="MALR"
29581..30086

REFERENCE 3 (bases 1 to 65335)
AUTHORS Waterston,R.H.
JOURNAL Direct Submission
Submitted (22-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 65335)
AUTHORS Waterston,R.H.
JOURNAL Direct Submission
Submitted (30-APR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 65335)
AUTHORS Waterston,R.H.
JOURNAL Direct Submission
Submitted (05-JUN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 6 (bases 1 to 65335)
AUTHORS Waterston,R.
JOURNAL Direct Submission
Submitted (26-JUN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jun 5, 2002 this sequence version replaced gi:20340530.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wustl.wustl.edu
----- Summary Statistics -----
Center project name: H_NH0692E14

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
Tateno,M., Catanesse,J.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.chori.org
VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-451F20, 2000 bp overlap;
the clone sequenced to the right is RP11-367N14. Actual start of
this clone is at base position 114362 of RP11-451F20.

There is a single plasmid region from 24260 to 24349.
Polymorphisms have been identified between AC107214 and AC112698.
Data from AC114911 was used to finish this clone, AC112698.
Location/Qualifiers
1. .65335
/organism="Homo sapiens"

/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-692E14"
/clone_lib="RPCI-11"
410. .571
/rpt_family="MIR"
1465. .1792
repeat_region
/rpt_family="Alu"
2272. .2563
repeat_region
/rpt_family="MAluR"
3341. .3644
repeat_region
/rpt_family="Alu"
3932. .4194
repeat_region
/rpt_family="MER2_type"
4196. .4356
repeat_region
/rpt_family="Alu"
4357. .4676
repeat_region
/rpt_family="Alu"
4689. .4739
repeat_region
/rpt_family="ERV1"
4757. .5034
repeat_region
/rpt_family="Alu"
5035. .5309
repeat_region
/rpt_family="L1"
5897. .6262
repeat_region
/rpt_family="CT-rich"
6264. .6334
repeat_region
/rpt_family="T-rich"
6824. .6937
repeat_region
/rpt_family="MIR"
7324. .7368
repeat_region
/rpt_family="(CA)n"
7861. .7893
repeat_region
/rpt_family="(TG)n"
8695. .8781
repeat_region
/rpt_family="MIR"
9092. .9128
repeat_region
/rpt_family="ERV1"
9129. .9435
repeat_region
/rpt_family="Alu"
9436. .9480
repeat_region
/rpt_family="ERV1"
9458. .9789
repeat_region
/rpt_family="ERV1"
10110. .10145
repeat_region
/rpt_family="L2"
10148. .10445
repeat_region
/rpt_family="Alu"
10793. .11080
repeat_region
/rpt_family="Alu"
12115. .12543
repeat_region
/rpt_family="L2"
12586. .12709
repeat_region
/rpt_family="Alu"
12720. .13003
repeat_region
/rpt_family="Alu"
13005. .13018
repeat_region
/rpt_family="(CA)n"
13019. .13053
repeat_region
/rpt_family="(CA)n"
13091. .13212
repeat_region
/rpt_family="Alu"
13245. .14096
repeat_region
/rpt_family="L2"
14285. .14596
repeat_region
/rpt_family="Alu"
16044. .16336
repeat_region
/rpt_family="Alu"
16463. .16779
repeat_region
/rpt_family="Alu"
16828. .16887
repeat_region

```
repeat_region      /rpt_family="MaLR"      16892..16976
repeat_region      /rpt_family="Alu"      17865..17961
repeat_region      /rpt_family="Cr-rich"    18410..18722
repeat_region      /rpt_family="Alu"      18947..19235
repeat_region      /rpt_family="Alu"      19442..19550
repeat_region      /rpt_family="MER1_type"    19936..20233
repeat_region      /rpt_family="Alu"      20486..20628
repeat_region      /rpt_family="MR"      21697..21959
repeat_region      /rpt_family="Alu"      22113..22156
repeat_region      /rpt_family="ERV1"      22197..22461
repeat_region      /rpt_family="Alu"      22482..22956
repeat_region      /rpt_family="ERV1"      23017..23315
repeat_region      /rpt_family="Alu"      23397..23533
repeat_region      /rpt_family="L2"      23623..23947
repeat_region      /rpt_family="MER2_type"    24060..24688
repeat_region      /rpt_family="L1"      24840..25268
repeat_region      /rpt_family="MaLR"      25269..25457
repeat_region      /rpt_family="L1"      25269..25457
```

Query Match	19.5%	Score 202.2;	DB 9;	Length 65335;
Best Local Similarity	78.1%;	Pred. No. 1.1e-29;		
Matches 243;	Conservative 0;	Mismatches 68;	Indels 0;	Gaps 0;

Qy	723	AGATACCTTACTTTTATGTCACAGGGGGCTGGGGCGCATATGCTGATGAGCCCTGTGCTCCCACTGCT	782
Db	50458	AGAAATTTAGTCTGAGAAACAGAGGCTGGGTGTGTGCTCAGCCTGTATATCCCAACACT	50399
Qy	783	TTGCGGGGCGCAGGCGAGGTGGATCACTTAAAGTCAGAGATTGTGAGACCAAGCCTGGCCAAC	842
Db	50398	TTGGGAGGCTGAGGCGAGGTGGATCACTGAGGTCAGAGATCGAGACCAAGTCTGGGCTTAG	50339
Qy	843	ATGGTGAACGGTGTCTCTCTATAAAATATCAAAATTTAGACAGGGCGGTGTGGCACACATC	902
Db	50338	ATGGTGAACCCCACTCTCTATTAATAATACAAATAATTAGCTGGGTGTGATGTGGCACCC	50279
Qy	903	TGTATATCCAGCTACTCAGGAGGCTAAACAGGAAATTTCTTTGAACCTGGAGGCGAAG	962
Db	50278	TGTATATCCAGCTACTTACGAGGCGCAAGGACGAGAGATGCTTGAACCCGGGAGGCGAGAT	50219
Qy	963	GTTGCACTGAGCACTTGTSCATCTCAGAGCCTGGGCAACAAGTGAGACTCTTGCTGTCAAAAAA	1022
Db	50218	GTTGCACTGAAACGCTGTCACTCAAGCCTGGGCAACATTAACAGATCTGTCTCAAAAAA	50159
Qy	1023	AAAAAAAAAAAA 1033	
Db	50158	AAAAAAAAAGA 50148	

Search completed: March 6, 2005, 05:42:31
Job time : 4571 secs

Job time : 4571 sec8

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2005, 04:05:58 ; Search time 600 Seconds
(Without alignments)
10241.153 Million cell updates/sec

Title: US-10-664-358-36

Perfect score: 1038
Sequence: 1 gnaatcgagcagcaggttaac.....aaaaaaaaaaacacgga 1038

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1035	99.7	1038	2	AAx51747 DNA encod
2	1035	99.7	1038	6	ABQ92599 Human sec
3	1035	99.7	1038	8	ACC50369 Human sec
4	1035	99.7	1038	8	ABZ71204 Human sec
5	1035	99.7	1038	9	ADB91088 Human sec
6	1035	99.7	1038	10	ADC73416 Human sec
7	200	19.3	68196	13	ABD32800 Human can
8	199.8	19.2	132748	6	ABT10719 Human bre
9	199.2	19.2	92500	12	ADP45591 Human int
10	199.2	19.2	110000	10	ACF42745 Human ALM
11	199.2	19.2	310268	13	ABD32548 Human can
12	198.2	19.1	32249	4	AA104931 Human tes
13	198.2	19.1	32249	4	ABL97825 Human rep
14	196.6	18.9	4165	9	AA161140 Human 76P
15	196.6	18.9	4165	12	ADM98105 Human nat
16	196.6	18.9	6837	9	AA161118 Human 76P
17	196.6	18.9	10394	12	ADM98083 Human nat
18	196.6	18.9	10394	12	AA161117 Human tum
19	196.6	18.9	10394	12	ADM98082 Human tum
20	196.4	18.9	4297	4	AA541953 Genomic B

ALIGNMENTS

RESULT 1	AAx51747	standard; DNA, 1038 BP.
ID	AAx51747	
XX	AAx51747	
AC	AAx51747	
DT	17-JUN-1999	(first entry)
XX		
DE	DNA encoding a human secreted protein.	
XX		
KW	Human secreted protein; cancer; immune disorder; infection;	
KW	inflammatory disorder; skin disorder; tumour; atherosclerosis;	
KW	reticulos; autoimmune disorder; Alzheimer's disease;	
KW	peripheral neuropathy; trauma; spinal cord injury; allergy;	
KW	hematopoietic disorder; skeletal disorder; neurological disorder;	
KW	arthritic disorder; asthma; immunodeficiency disease; AIDS;	
KW	transplant rejection; ss.	
XX		
OS	Homo sapiens.	
XX		
FN	W09911293-AL.	
XX		
PD	11-MAR-1999.	
XX		
PF	03-SEP-1998;	98WC-US018360.
XX		
PR	05-SEP-1997;	97US-0057626P.
XX		
PR	05-SEP-1997;	97US-0057663P.
XX		
PR	12-SEP-1997;	97US-0058666P.
XX		
PR	12-SEP-1997;	97US-0058667P.
XX		
PR	12-SEP-1997;	97US-0058973P.
XX		
PR	12-SEP-1997;	97US-0058974P.
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Moore PA, Ruben SM, Lafleur DW, Shi Y, Rosen GA, Olsen HS;	
XX	Ehner R, Brewer LA;	
XX	WPI; 1999-204988/17.	
DR	P-PSDB; AA12960.	
XX		
PT	New isolated human genes and the secreted polypeptides they encode -	

PT useful for diagnosis and treatment of e.g. neurological disorders,
PT tumors, immune disorders, inflammation or hematological disorders.
XX
PS Claim 1, Page 176-177, 215pp; English.
XX
CC AAX51701-55 encode human secreted proteins. The polynucleotides and their
CC corresponding secreted polypeptides are useful for preventing, treating
CC or ameliorating medical conditions, e.g. by protein or gene therapy.
CC Pathological conditions can also be diagnosed by determining the amount
CC of the new polypeptides in a sample or by determining the presence of
CC mutations in the new polynucleotides. Specific uses are described for
CC each polynucleotide, based on which tissues they are most highly
CC expressed in, and include developing products for the diagnosis or
CC treatment of cancer, immune disorders, infection, inflammatory disorders,
CC skin disorders, tumours, atherosclerosis, restenosis, autoimmune
CC disorders, Alzheimer's disease, peripheral neuropathies, trauma, spinal
CC cord injuries, allergy, hematopoietic disorders, skeletal disorders,
CC neurological disorders, arthritic disorders, asthma, immunodeficiency
CC diseases, AIDS and transplant rejection. The polypeptides are also useful
CC for identifying their binding partners
XX
SQ Sequence 1038 BP; 328 A; 167 C; 198 G; 339 T; 0 U; 6 Other;
Query Match 99.7%; Score 1035; DB 2; Length 1038;
Best Local Similarity 100.0%; Pred. No. 4.2e-213;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GNAATTGGCAGAGTTAATGTAATAATTTCTATATGAATTTTAATGGAAATTGA 60
DB 1 GNAATTGGCAGAGTTAATGTAATAATTTCTATATGAATTTTAATGGAAATTGA 60
QY 61 GCATCATGAAAAAATGCTTACTGTTGAAAAAATTTTGTATACATTTTGGTCAACTA 120
DB 61 GCATCATGAAAAAATGCTTACTGTTGAAAAAATTTTGTATACATTTTGGTCAACTA 120
QY 121 ATCTTCAATACTTTTATGTAATTAATTTTAAAGTTTACAGAGTCTTATATAG 180
DB 121 ATCTTCAATACTTTTATGTAATTAATTTTAAAGTTTACAGAGTCTTATATAG 180
QY 181 TAAATGGCAGCTGACAGATGAAATATACATATCAATTTTGTACTATCTTATAGG 240
DB 181 TAAATGGCAGCTGACAGATGAAATATACATATCAATTTTGTACTATCTTATAGG 240
QY 241 AAAATCAGAGAATTTCAAAACCTTGTATGTTTAAAGGTATAGTCACTTTTATTAATGT 300
DB 241 AAAATCAGAGAATTTCAAAACCTTGTATGTTTAAAGGTATAGTCACTTTTATTAATGT 300
QY 301 GCGGTATTTTATACATGATTTGACGTTTGTGMAAATTTTCCCTGGACTTTTATTTTA 360
DB 301 GCGGTATTTTATACATGATTTGACGTTTGTGMAAATTTTCCCTGGACTTTTATTTTA 360
QY 361 GATGAGATCTACAGTGTAGCAAACTTATATCTGCAACCTGATTAAGTGCATAGTC 420
DB 361 GATGAGATCTACAGTGTAGCAAACTTATATCTGCAACCTGATTAAGTGCATAGTC 420
QY 421 AGACTCATCCCATGCTTAAATTAATTAAGTTGKAAATACGTTTGTAAATAGTTGTGT 480
DB 421 AGACTCATCCCATGCTTAAATTAATTAAGTTGKAAATACGTTTGTAAATAGTTGTGT 480
QY 481 AGGTCAATTAACAGAGCTTCAAGGKATTAATTAATTAATTAATTAATTAATTAAT 540
DB 481 AGGTCAATTAACAGAGCTTCAAGGKATTAATTAATTAATTAATTAATTAATTAAT 540
QY 541 TGAATACCGTTTTCATGCAAAAGTAAATCTTCAGCCTTTATTTTATTTATTAAT 600
DB 541 TGAATACCGTTTTCATGCAAAAGTAAATCTTCAGCCTTTATTTTATTTATTAAT 600
QY 601 ATATTAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
DB 601 ATATTAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 661 GCTTTGCTAAATTTTAACTGACAAATCTTACGAGTCTTCAATTTTGTGAGAT 720

DB 661 GCTTTGCTAAATTTTAACTGACAAATCTTACGAGTCTTCAATTTTGTGAGAT 720
QY 721 GAAGATATCTTATGTTTATGTCAGAGGGCTGGGCGGATAGCTGATGCTGTGTCAGTG 780
DB 721 GAAGATATCTTATGTTTATGTCAGAGGGCTGGGCGGATAGCTGATGCTGTGTCAGTG 780
QY 781 CTTTGGCGGGCCGAGGAGGTGATTCATTAAAGTTCAGAGTTTGAACAGCCCTGCGCA 840
DB 781 CTTTGGCGGGCCGAGGAGGTGATTCATTAAAGTTCAGAGTTTGAACAGCCCTGCGCA 840
QY 841 ACATGTGAAACGTTTCTCTACTCAATTAATAATAATTAAGACAGGGGTGTGGCACACA 900
DB 841 ACATGTGAAACGTTTCTCTACTCAATTAATAATAATTAAGACAGGGGTGTGGCACACA 900
QY 901 TCTGTAAATTCAGTACTTACAGAGGCTTAAACAGAGAAATTCCTTGAACCTGGAGAGCAG 960
DB 901 TCTGTAAATTCAGTACTTACAGAGGCTTAAACAGAGAAATTCCTTGAACCTGGAGAGCAG 960
QY 961 AGGTTCAGAGGAGCATTGACATCCAGCTGGGCAACAGGTGAGACTTGTCTCAAAA 1020
DB 961 AGGTTCAGAGGAGCATTGACATCCAGCTGGGCAACAGGTGAGACTTGTCTCAAAA 1020
QY 1021 AAAAAAAAAAACTCGA 1038
DB 1021 AAAAAAAAAAACTCGA 1038
RESULT 2
AB092599
ID AB092599 standard; cDNA; 1038 BP.
XX
AC AB092599;
XX
DT 12-NOV-2002 (first entry)
XX
DE Human secreted protein encoding cDNA SEQ ID NO 57.
XX
KW Human; noctropic; neuroprotective; cytosolic; dermatological; virucide;
KW immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; vulnary;
KW antiparkinsonian; antistroke; antianemic; antiarthritic; cancer;
KW antithrombotic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antilucer; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine;
KW gene; ss.
XX
OS Homo sapiens.
XX
PN WO200257420-A2.
XX
PN 25-JUL-2002.
XX
PF 17-JAN-2002; 2002MO-US001109.
XX
PR 18-JAN-2001; 2001US-0262066P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Moore PA, Ruben SM, Lafleur DW, Shi Y, Rosen CA, Olsen H;
PI Ebner R, Brewer LA;
XX
DR P-PSDB; ABP62059.
XX
DR WPI; 2002-599716/64.
XX
PT New polynucleotides and polypeptides useful for diagnosing, prognosing,
PT treating or preventing e.g. neurodegenerative, central nervous system,
PT autoimmune, respiratory, reproductive, or inflammatory diseases or
PT disorders.
XX
PS Claim 1, Page 726; 785pp; English.
XX
CC The invention relates to novel genes (AB092553-AB092607) and proteins
CC (ABP62013-ABP62153) useful for preventing, treating or ameliorating

CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (anti)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections
XX

XX Sequence 1038 BP; 328 A; 167 C; 198 G; 339 T; 0 U; 6 Other;

Query Match 99.7%; Score 1035; DB 6; Length 1038;
Best Local Similarity 100.0%; Pred. No. 4.2e-213;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 GNAATGGGCGACGATTAATGTAATAATTTCTATATGAATTTAATGGGAATTGA 60
Db 1 GNAATGGGCGACGATTAATGTAATAATTTCTATATGAATTTAATGGGAATTGA 60
Qy 61 GCATCATAGAAAAAATGCTCTTACTGTGAAAAATTATTTGTACATTTTGTCAACTA 120
Db 61 GCATCATAGAAAAAATGCTCTTACTGTGAAAAATTATTTGTACATTTTGTCAACTA 120
Qy 121 ATCTTCAATACCTTTAGTAAATGTTAAGTTGACAGTGGCAGCTTTATATAG 180
Db 121 ATCTTCAATACCTTTAGTAAATGTTAAGTTGACAGTGGCAGCTTTATATAG 180
Qy 181 TAAATGCGAGTGCACAGTAATAAATACATATATTTTGTGACTTCTTATAGG 240
Db 181 TAAATGCGAGTGCACAGTAATAAATACATATATTTTGTGACTTCTTATAGG 240
Qy 241 AAAATCAGAGATTTCAAAAATCTTGTAGTTTGTAGGTAATGACATTTTATATAATGT 300
Db 241 AAAATCAGAGATTTCAAAAATCTTGTAGTTTGTAGGTAATGACATTTTATATAATGT 300
Qy 301 GCGGTAATTTATACATGATTTGACGTTTGTMAAATATTTTCCGTGACATTTTATTTTA 360
Db 301 GCGGTAATTTATACATGATTTGACGTTTGTMAAATATTTTCCGTGACATTTTATTTTA 360
Qy 361 GATAGATCTACAGTGTAGGCAAACTTAATATCTGTCACTCATTAAGTGTATGTC 420
Db 361 GATAGATCTACAGTGTAGGCAAACTTAATATCTGTCACTCATTAAGTGTATGTC 420
Qy 421 AGACTCATCCCATGCTTAAATTTAGTGTGKAAATACGCTTTGTAAATAGTTGTT 480
Db 421 AGACTCATCCCATGCTTAAATTTAGTGTGKAAATACGCTTTGTAAATAGTTGTT 480
Qy 481 AGGTCATTAATCACCAAGCTTTCAAGKATTAATTATATAAAACCTGGKTTTATTCCTG 540
Db 481 AGGTCATTAATCACCAAGCTTTCAAGKATTAATTATATAAAACCTGGKTTTATTCCTG 540
Qy 541 TGAATATCCGTTTTTTCATGCAAAAGTTAAATCTTCAGCTTTAATTTTTTATATAT 600
Db 541 TGAATATCCGTTTTTTCATGCAAAAGTTAAATCTTCAGCTTTAATTTTTTATATAT 600
Qy 601 ATATTAAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Db 601 ATATTAAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Qy 661 GCTTTTCTAAATTTGTTACCTGCAAAAATCTTACGCAATTTTCTTATTTTGGAT 720
Db 661 GCTTTTCTAAATTTGTTACCTGCAAAAATCTTACGCAATTTTCTTATTTTGGAT 720
Qy 721 GAAGATCTTAAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Db 721 GAAGATCTTAAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Qy 781 CTTTGGGGGCGAGGAGGTGATCACTTAAAGTTCAGAGATTGAGACACAGCTGACCA 840
```

```
Db 781 CTTTGGGGGCGAGGAGGTGATCACTTAAAGTTCAGAGATTGAGACACAGCTGACCA 840
Qy 841 ACATGTGAAACGTTGTCTTACTTAAATAATTCANAAATTTAGACAGCGTGTGCAACA 900
Db 841 ACATGTGAAACGTTGTCTTACTTAAATAATTCANAAATTTAGACAGCGTGTGCAACA 900
Qy 901 TCTGTATTCAGCTTACTCAGAGGCTTACACAGAAATTCCTTGAACCTGGAGGAG 960
Db 901 TCTGTATTCAGCTTACTCAGAGGCTTACACAGAAATTTCTTGAACCTGGAGGAG 960
Qy 961 AGGTTGAGTGAAGCATTTGCACTGCAAGCTGGGCAACAGTGAAGCTTTGTCAAAA 1020
Db 961 AGGTTGAGTGAAGCATTTGCACTGCAAGCTGGGCAACAGTGAAGCTTTGTCAAAA 1020
Qy 1021 AAAAAAAAAAAATCTGCA 1038
Db 1021 AAAAAAAAAAAATCTGCA 1038
```

RESULT 3

ACC50369 standard; cDNA; 1038 BP.
ACC50369;
12-JUN-2003 (first entry)
XX
DE Human secreted protein coding sequence, SEQ ID 36.
XX
KW Cardiac; antiarrhythmic; antiarteriosclerotic; vasotropic; cyostatic;
KW vulnerable; antiinflammatory; nootropic; neuroprotective;
KW antiparkinsonian; gene therapy; human; cardiovascular disorder; gene; ss.
XX Homo sapiens.
XX
PN MO200295010-A2.
XX
PD 28-NOV-2002.
XX
PF 19-MAR-2002; 2002MO-US009785.
XX
PR 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2003-129429/12.
XX
PT Novel human secreted proteins, useful for detecting, preventing,
PT diagnosing, prognosticating, treating and/or ameliorating cardiovascular
PT disorders such as arrhythmia.
XX
PS Claim 21; SEQ ID NO 36; 1881bp; English.
XX
The present invention relates to novel human secreted proteins (ABR4/633-
ABR48145) and their coding sequences (ACC50344-ACC50856). The proteins
and their coding sequences are useful for the preparation of a diagnostic
or pharmaceutical composition for diagnosing or treating a cardiovascular
disorder (e.g., arrhythmia, tachycardia, cardiac arrest, coronary
arteriosclerosis and myocardial ischaemia), neural disorders, immune
system disorders, muscular disorders, reproductive disorders,
gastrointestinal disorders, pulmonary disorders, renal disorders,
proliferative disorders and/or cancerous diseases and conditions, for
wound healing and epithelial cell proliferation, to treat inflammation or
infection, for treating thrombosis and arteriosclerosis, for treating or
preventing neural damage which occurs in neuronal disorders or
neurodegenerative conditions such as Alzheimer's disease and Parkinson's
disease, to enhance bone and periodontal regeneration and aid in tissue
transplants or bone grafts, to prevent skin aging or hair loss, to

CC stimulate growth and differentiation of haematopoietic cells and bone
CC marrow cells when used in combination with other cytokines, to maintain
CC organs before transplantation or for supporting cell culture of primary
CC tissues, to increase or decrease differentiation or proliferation of
CC embryonic stem cells, or to modulate mammalian characteristics or
CC metabolism. Note: The sequence data for this patent was published in
CC electronic format and is available from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

Sequence 1038 BP; 328 A; 167 C; 198 G; 339 T; 0 U; 6 Other;

Query Match 99.7%; Score 1035; DB 8; Length 1038;
Best Local Similarity 100.0%; Pred. No. 4.2e-213;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GNAATTCGGACGACGATTAAGTAAATTTCTTAATGAATTTTAATGGGAATTAGA 60
DB 1 GNAATTCGGACGACGATTAAGTAAATTTCTTAATGAATTTTAATGGGAATTAGA 60
QY 61 GCATCATAGAAAAAATGCTCTTACTGTGAAAACATTATTTGTTACATTTTGCTCAACTA 120
DB 61 GCATCATAGAAAAAATGCTCTTACTGTGAAAACATTATTTGTTACATTTTGCTCAACTA 120
QY 121 ATCTTTCATTAATTTTACTTAATTAATGTTAAGTTGACAGAGGAGCTTATATAG 180
DB 121 ATCTTTCATTAATTTTACTTAATTAATGTTAAGTTGACAGAGGAGCTTATATAG 180
QY 181 TAAATGGACGCTGACGATGAAATTAATTAATTTTGTGACATCTTATTTAG 240
DB 181 TAAATGGACGCTGACGATGAAATTAATTAATTTTGTGACATCTTATTTAG 240
QY 241 AAAATCAGAAATTTCAAAACCTTTGTTAGTTTGAAGTATAGTACATTTTAATAGT 300
DB 241 AAAATCAGAAATTTCAAAACCTTTGTTAGTTTGAAGTATAGTACATTTTAATAGT 300
QY 301 GCGGTATATTTATACATGATTTGAGCTTGTGMAAATTTTCCCTGGAATTTATTTTA 360
DB 301 GCGGTATATTTATACATGATTTGAGCTTGTGMAAATTTTCCCTGGAATTTATTTTA 360
QY 361 GATGAGATCTACAGTGTAGGCAAACTTAATTAATCTGCACTCCATAGGTGATATGTC 420
DB 361 GATGAGATCTACAGTGTAGGCAAACTTAATTAATCTGCACTCCATAGGTGATATGTC 420
QY 421 AGACATCAATCCCATGCTTAAATTAATTAATGTTGAAAATACGTTTGTAAATAGTTGTT 480
DB 421 AGACATCAATCCCATGCTTAAATTAATTAATGTTGAAAATACGTTTGTAAATAGTTGTT 480
QY 481 AGGTCAATATCAACAAGCTTCAAGGKATTAATTAATTAATTAATTAATTAATTAAT 540
DB 481 AGGTCAATATCAACAAGCTTCAAGGKATTAATTAATTAATTAATTAATTAATTAAT 540
QY 541 TGAATATMCCGTTTTTTCATGCAAAAGTTAAATTTCTCAGCCCTTAATTTTATTAAT 600
DB 541 TGAATATMCCGTTTTTTCATGCAAAAGTTAAATTTCTCAGCCCTTAATTTTATTAAT 600
QY 601 ATATAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 ATATAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 GCTTTTGTAAATTTTACTGACAAATCTTAAGCCAGTTCTTCAATTTTGTTTGAAT 720
DB 661 GCTTTTGTAAATTTTACTGACAAATCTTAAGCCAGTTCTTCAATTTTGTTTGAAT 720
QY 721 GAAATATCTTAATTTTGTGACGAGGCTGAGGCGGATGATGATGATGATGATGATGAT 780
DB 721 GAAATATCTTAATTTTGTGACGAGGCTGAGGCGGATGATGATGATGATGATGATGAT 780
QY 781 CTTTGGCGGCGGACGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 CTTTGGCGGCGGACGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 ACATGTGAAACGTTGTTCTTACTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
DB 841 ACATGTGAAACGTTGTTCTTACTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900

```

```

DB 841 ACATGTGAAACGTTGTTCTTACTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
QY 901 TCTGTAATTTCCAGGCTACTCAGAGGCTTAACAACAGAAAATTTCTTGAACCTGGAGGACAG 960
DB 901 TCTGTAATTTCCAGGCTACTCAGAGGCTTAACAACAGAAAATTTCTTGAACCTGGAGGACAG 960
QY 961 AGTTGACGTGAGCCATTGCACTTCAGCCTGGGCAACAGTGAAGACTTGTCTCAAAA 1020
DB 961 AGTTGACGTGAGCCATTGCACTTCAGCCTGGGCAACAGTGAAGACTTGTCTCAAAA 1020
QY 1021 AAAAAAAAAAACTGCA 1038
DB 1021 AAAAAAAAAAACTGCA 1038

RESULT 4
AB271204
ID AB271204 standard; cDNA; 1038 BP.
XX
XX AC AB271204;
XX
XX 03-APR-2003 (first entry)
XX
XX Human secreted protein-encoding gene 15 cDNA clone HB1AE26, SEQ ID NO:25.
XX
XX Human; secreted protein; digestive disorder; gastrointestinal disorder;
XX mouth; oesophagus; stomach; small intestine; large intestine; liver;
XX biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
XX immune disorder; inflammation; infection; wound healing; drug screening;
XX chromosome identification; chromosome mapping; cytostatic;
XX antiinflammatory; immunosuppressive; vulnerability; gene therapy; gene, ss.
XX
XX Homo sapiens.
XX
XX WO200276488-A1.
XX
XX 03-OCT-2002.
XX
XX 19-MAR-2002; 2002WO-US008276.
XX
XX 21-MAR-2001; 2001US-0277340P.
XX 19-JUL-2001; 2001US-0306171P.
XX 13-NOV-2001; 2001US-0331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2003-029900/02.
XX P-PSDB; ABR00025.
XX
XX New human secreted proteins and nucleic acids, useful for detecting,
XX preventing, diagnosing, prognosticating, treating and/or ameliorating
XX e.g. gastrointestinal diseases and disorders, or cancers.
XX
XX Claim 21; Page 764-765; 1216pp; English.
XX
XX AB271190-AB271478 represent cDNAs corresponding to 178 human secreted
XX protein genes, and ABP00011-ABP00299 represent the proteins they encode.
XX AB271479-AB271540 represent human secreted protein genomic fragments. The
XX invention also encompasses antibodies specific for the secreted proteins,
XX the use of the secreted proteins in drug screening, and recombinant
XX vectors and host cells comprising a nucleic acid of the invention. The
XX secreted proteins, nucleic acids encoding them, antibodies or antibody
XX fragments specific for the secreted proteins, treating, and modulators of protein
XX activity are useful for diagnosing, treating, and modulators of preventing
XX digestive disorders. Such conditions include disorders of the mouth,
XX oesophagus, stomach, small intestine, large intestine, liver, biliary
XX tract and pancreas, and include cancers of these organs and tissues. The
XX secreted proteins and their nucleic acids may also be used in the
XX treatment of immune disorders, inflammation, infection,
XX hyperproliferative disorders, and to promote wound healing. Nucleic acids
XX of the invention may be used for chromosome identification, chromosome

```


CC mapping, in gene therapy, for identifying individuals from minute
 CC biological samples, as hybridisation probes, and as molecular weight
 CC markers. The present sequence represents a human secreted protein-
 CC encoding cDNA clone of the invention

XX Sequence 1038 BP; 328 A; 167 C; 198 G; 339 T; 0 U; 6 Other;

Query Match 99.7%; Score 1035; DB 8; Length 1030;
 Best Local Similarity 100.0%; Pred. No. 4.2e-213;
 Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GNAATTGGCGACGAGTTAATGATATAAATTTCTATATGAATTTTAATGGGAATTAGA 60
DB 1 GNAATTGGCGACGAGTTAATGATATAAATTTCTATATGAATTTTAATGGGAATTAGA 60
QY 61 GCATCATAGAAAAAATGCTCTTACTGTGGAAAAACATTAATTTGTATACATTTGGTCAACTA 120
DB 61 GCATCATAGAAAAAATGCTCTTACTGTGGAAAAACATTAATTTGTATACATTTGGTCAACTA 120
QY 121 ATCTTTCAATTAATCTTTAGTAACTATATAATGTTAAGTTGACAGTGGCAGTCTTATATAG 180
DB 121 ATCTTTCAATTAATCTTTAGTAACTATATAATGTTAAGTTGACAGTGGCAGTCTTATATAG 180
QY 181 TAAATGGCAGCTGACAGCATGAAAAATACATATCTAATATTTTGTGACTATCTTATAGG 240
DB 181 TAAATGGCAGCTGACAGCATGAAAAATACATATCTAATATTTTGTGACTATCTTATAGG 240
QY 241 AAAATCAGAGAAATTTCAAAAACCTGTTAGTTTAAAGGTATAGTCACTTTTATTAATGT 300
DB 241 AAAATCAGAGAAATTTCAAAAACCTGTTAGTTTAAAGGTATAGTCACTTTTATTAATGT 300
QY 301 GCGGTATATTATATCATGATTTGACGTTTGMAAATTTTCCCTGGACCTTATATTGA 360
DB 301 GCGGTATATTATATCATGATTTGACGTTTGMAAATTTTCCCTGGACCTTATATTGA 360
QY 361 GATGAGATCTACAGTGTAGGCAAACTATATATCTGTCAACCTCATTTAGTGTATAGTC 420
DB 361 GATGAGATCTACAGTGTAGGCAAACTATATATCTGTCAACCTCATTTAGTGTATAGTC 420
QY 421 AGACTCATCCCGCAGCTTAAATTTATAGTTGTAAAAAATACGTTTGTAAATAGTTGTG 480
DB 421 AGACTCATCCCGCAGCTTAAATTTATAGTTGTAAAAAATACGTTTGTAAATAGTTGTG 480
QY 481 AGGTCATTATCAGCAAGCTTCAAGGKATTAATTATTAATAAACCCTGGKTTTATTCCTG 540
DB 481 AGGTCATTATCAGCAAGCTTCAAGGKATTAATTATTAATAAACCCTGGKTTTATTCCTG 540
QY 541 TGAATAMCCGTTTTTTCATGCAAAAGTTAAATTTCTTCAAGCCTTTAATTTTATTTAAT 600
DB 541 TGAATAMCCGTTTTTTCATGCAAAAGTTAAATTTCTTCAAGCCTTTAATTTTATTTAAT 600
QY 601 ATATTAAGAGATGTGATGATAGTACTACAAAACAGAAAAAATAACAGATTTGCTTTGTG 660
DB 601 ATATTAAGAGATGTGATGATAGTACTACAAAACAGAAAAAATAACAGATTTGCTTTGTG 660
QY 661 GCTTTTCTAAATTTGTTAATCTGACAAAATCTTACCCAGTTTCTTATTTTCTTTTGAT 720
DB 661 GCTTTTCTAAATTTGTTAATCTGACAAAATCTTACCCAGTTTCTTATTTTCTTTTGAT 720
QY 721 GAAGATATTAGTTTATAGTCCAGGGGCTGGCGGAGATAGCTGATGCTGTGCTCCAGTG 780
DB 721 GAAGATATTAGTTTATAGTCCAGGGGCTGGCGGAGATAGCTGATGCTGTGCTCCAGTG 780
QY 781 CTTTGGCGGGCCGAGGCGAGTGTATCACTTAAGGTCAAGAGTTTGAAGACAGCTGCCCA 840
DB 781 CTTTGGCGGGCCGAGGCGAGTGTATCACTTAAGGTCAAGAGTTTGAAGACAGCTGCCCA 840
QY 841 ACATGTGGAAGCGTTGCTCTACTATTAATAATCAAAATTTTGAACAGGGGTGTGGCACACA 900
DB 841 ACATGTGGAAGCGTTGCTCTACTATTAATAATCAAAATTTTGAACAGGGGTGTGGCACACA 900
QY 901 TCTGTATTTCCAGCTACTCAGAGAGGCTTAACAGAAAAATTTCTTGAACCTGGAGGCGAG 960
DB 901 TCTGTATTTCCAGCTACTCAGAGAGGCTTAACAGAAAAATTTCTTGAACCTGGAGGCGAG 960

```

```

DB 901 TCTGTATTTCCAGCTACTCAGAGAGGCTTAACAGAAAAATTTCTTGAACCTGGAGGCGAG 960
QY 961 AGGTTGACGTAGAGCATTTGCACTCCAGCTGGGCAACAGATGAGACTTGTCTCAAAA 1020
DB 961 AGGTTGACGTAGAGCATTTGCACTCCAGCTGGGCAACAGATGAGACTTGTCTCAAAA 1020
QY 1021 AAAAAAAAAAAAACTCGA 1038
DB 1021 AAAAAAAAAAAAACTCGA 1038

```

RESULT 5

ADB91088
 ID ADB91088 standard; cDNA; 1038 BP.

AC ADB91088;

DT 04-DEC-2003 (first entry)

DE Human secreted protein cDNA #SEQ ID 34.

KM Secreted protein; gene therapy; antidiabetic; diabetes; human; ss.

OS Homo sapiens.

PN MO2003004622-A2.

PD 16-JAN-2003.

PE 19-MAR-2002; 2002MO-US008124.

PR 21-MAR-2001; 2001US-0277340P.

PR 19-JUL-2001; 2001US-0306171P.

PR 13-NOV-2001; 2001US-0331287P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

XX WPI; 2003-229407/22.

XX Nucleic acid encoding a human secreted protein is useful in diagnosing or

XX treating diabetes or conditions related to diabetes.

XX Claim 9; SEQ ID NO 34; 1537bp; English.

XX The invention relates to isolated nucleic acid molecules ADB91065-

XX ADB91448 and ADB91835-ADB91911 encoding human secreted proteins ADB91449-

XX ADB91834. Also disclosed is a recombinant vector comprising a

XX polynucleotide of the invention, and a recombinant host cell comprising

XX the recombinant vector. The polypeptide of the invention is useful in

XX identifying a binding partner by contacting the polypeptide with a

XX binding partner, and determining whether the binding partner increases or

XX decreases activity of the polypeptide. The polypeptide, polynucleotide,

XX antibody or its fragment, agonist or antagonist are useful for preparing

XX a pharmaceutical composition for diagnosing or treating diabetes or

XX immunoglobulin Fc portion used to generate fusion proteins, increasing

XX the stability of the fused protein as compared to the secreted protein

XX only. Note: The sequence data for this patent did not form part of the

XX printed specification, but was obtained in electronic format directly

XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1038 BP; 328 A; 167 C; 198 G; 339 T; 0 U; 6 Other;
 Query Match 99.7%; Score 1035; DB 9; Length 1038;
 Best Local Similarity 100.0%; Pred. No. 4.2e-213;
 Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GNAATTGGCGACGAGTTAATGATATAAATTTCTATATGAATTTTAATGGGAATTAGA 60
DB 1 GNAATTGGCGACGAGTTAATGATATAAATTTCTATATGAATTTTAATGGGAATTAGA 60

```

QY	6	GCATCATGAAAAAATGCTTACTGCTGTAAGAAACATATTTGGTATCATTTTGGTCAACTA	120
Db	61	GCATCATGAAAAAATGCTTACTGCTGTAAGAAACATATTTGGTATCATTTTGGTCAACTA	120
QY	121	ATCTTTCATATACTTTTAGTACTATAATGTTAACTGTACAGTGGCAGTCTTATATAG	180
Db	121	ATCTTTCATATACTTTTAGTACTATAATGTTAACTGTACAGTGGCAGTCTTATATAG	180
QY	181	TAAATGGAGCGTACAGAGATGAAAATTAACATATCTTAATTTTGGACTATCTTATAGG	240
Db	181	TAAATGGAGCGTACAGAGATGAAAATTAACATATCTTAATTTTGGACTATCTTATAGG	240
QY	241	AAATCAGAGATTTCAAAACCTGTGTAGTTTTAGGGTATAGTCACATTTTATAAATGT	300
Db	241	AAATCAGAGATTTCAAAACCTGTGTAGTTTTAGGGTATAGTCACATTTTATAAATGT	300
QY	301	GCGGTATATTTATACATGATTTGACAGTTTGGMAAATATTTTCCCTGACCTTTATTTTA	360
Db	301	GCGGTATATTTATACATGATTTGACAGTTTGGMAAATATTTTCCCTGACCTTTATTTTA	360
QY	361	GATGAGATCTACAGGTAGGCAAACTTATATATCTGCAACTCATTAGTGCATATGTC	420
Db	361	GATGAGATCTACAGGTAGGCAAACTTATATATCTGCAACTCATTAGTGCATATGTC	420
QY	421	AGACTCATCCCGACTAAAAATTAATAGTTGTAAATAACGCTTTTGTAAATAGTTGTGT	480
Db	421	AGACTCATCCCGACTAAAAATTAATAGTTGTAAATAACGCTTTTGTAAATAGTTGTGT	480
QY	481	AGGTCATATACCAAGTCTTCAAGGKATTCATATATATAAAACCTTGKTTTATCTGTG	540
Db	481	AGGTCATATACCAAGTCTTCAAGGKATTCATATATATAAAACCTTGKTTTATCTGTG	540
QY	541	TGAATAMCCGTTTTTTCATGCAAGTAAGTTAAATTTCTTCAGCCCTTATTTTATATAT	600
Db	541	TGAATAMCCGTTTTTTCATGCAAGTAAGTTAAATTTCTTCAGCCCTTATTTTATATAT	600
QY	601	ATATATAGATGTGATGATATGATCTACAAAACAGAAAAAATAAACAATTTCCGTTGTG	660
Db	601	ATATATAGATGTGATGATATGATCTACAAAACAGAAAAAATAAACAATTTCCGTTGTG	660
QY	661	GCTTTTGTAAATTTGTACTGACAAAATCTTACGCAAGTCTTCAATTTTCGTTTGAAGAT	720
Db	661	GCTTTTGTAAATTTGTACTGACAAAATCTTACGCAAGTCTTCAATTTTCGTTTGAAGAT	720
QY	721	GAAAGTACTTATGTTTATGTCACAGGGGCTGGGCGCATATGCTGATCCCTGTGCTCCAGTGT	780
Db	721	GAAAGTACTTATGTTTATGTCACAGGGGCTGGGCGCATATGCTGATCCCTGTGCTCCAGTGT	780
QY	781	CTTTCGGGGGCGGAGGAGGTGGATCACTTAAGTCAAGAGTTTATAGACAAGCTGGCCCA	840
Db	781	CTTTCGGGGGCGGAGGAGGTGGATCACTTAAGTCAAGAGTTTATAGACAAGCTGGCCCA	840
QY	841	ACATGATGAAACGTTGTCTCTACTAAAAATACAAAAATTAGACAAGCTGGTGACACACA	900
Db	841	ACATGATGAAACGTTGTCTCTACTAAAAATACAAAAATTAGACAAGCTGGTGACACACA	900
QY	901	TCTGTAAATTCGAGTACTCAGAGAGGCTTAACAAGAAAAATTCCTGAACTCGGGAGGGAG	960
Db	901	TCTGTAAATTCGAGTACTCAGAGAGGCTTAACAAGAAAAATTCCTGAACTCGGGAGGGAG	960
QY	961	AGGTTGACAGTGAAGCATTTGCACTCCAGCCTGGGCAACAAGTGAACCTTGTGTCAAAA	1020
Db	961	AGGTTGACAGTGAAGCATTTGCACTCCAGCCTGGGCAACAAGTGAACCTTGTGTCAAAA	1020
QY	1021	AAAAAAAAAAAAAATCTCGA 1038	
Db	1021	AAAAAAAAAAAAAATCTCGA 1038	

AC	ADC73416;	
AD		
AE	01-JAN-2004 (first entry)	
AF		
AG	Human secreted protein-related DNA - SEQ ID 49.	
AH		
AI	antianemic; antirheumatic; antiarthritic; antiinflammatory; antihypoid;	
AJ	antidiabetic; immunosuppressive; dermatological; nephrotoxic;	
AK	antiparkinsonian; neuroprotective; neurotropic; antibacterial; virocidic;	
AL	fungicide; antiparasitic; antiarteriosclerotic; vulnery; cytostatic;	
AM	haemopoietic; haematologic; anaemia; autoimmune disorder;	
AN	rheumatoid arthritis; inflammation; Grave's disease; diabetes;	
AO	systemic lupus erythematosus; glomerulonephritis; neurodegenerative;	
AP	Parkinson's; Alzheimer's; wound; hyperproliferative; atherosclerosis;	
AQ	cancer; bacterial; viral; fungal; parasitic infection; gene therapy;	
AR	human; gene; ds.	
AS		
AT	Homo sapiens.	
AW		
AX	WO2003038063-A2.	
AY		
AZ	08-MAY-2003.	
BA		
BB	19-MAR-2002; 2002WO-US008277.	
BC		
BD	21-MAR-2001; 2001US-0277340P.	
BE	PR 19-JUL-2001; 2001US-0306171P.	
BF	PR 13-NOV-2001; 2001US-0331287P.	
BG		
BH	(HUMA-) HUMAN GENOME SCI INC.	
BI		
BJ	Rosen CA, Ruben SM;	
BK	WPI; 2003-430516/40.	
BL	P-PSDB; ADC74031.	
BM		
BN	New human secreted polypeptide for diagnosing, preventing or treating	
BO	hematopoietic or hematologic disorders (e.g. anemia), autoimmune	
BP	disorders (e.g. diabetes) or hyperproliferative disorders (e.g. cancer or	
BQ	atherosclerosis).	
BR		
BS	Claim 27; SEQ ID NO 49; 2272pp; English.	
BT		
BU		
BV		
BW		
BX		
BY		
BZ		
CA		
CB		
CC		
CD		
CE		
CF		
CG		
CH		
CI		
CJ		
CK		
CL		
CM		
CN		
CO		
CP		
CQ		
CR		
CS		
CT		
CU		
CV		
CW		
CX		
CY		
CA		
CB		
CC		
CD		
CE		
CF		
CG		
CH		
CI		
CJ		
CK		
CL		
CM		
CN		
CO		
CP		
CQ		
CR		
CS		
CT		
CU		
CV		
CW		
CX		
CY		
CA		
CB		
CC		
CD		
CE		
CF		
CG		
CH		
CI		
CJ		
CK		
CL		
CM		
CN		
CO		
CP		
CQ		
CR		
CS		
CT		
CU		
CV		
CW		
CX		
CY		
CA		
CB		
CC		
CD		
CE		
CF		
CG		
CH		
CI		
CJ		
CK		
CL		
CM		
CN		
CO		
CP		
CQ		
CR		
CS		
CT		
CU		
CV		
CW		
CX		
CY		
CA		
CB		
CC		
CD		
CE		
CF		
CG		
CH		
CI		
CJ		
CK		
CL		
CM		
CN		
CO		
CP		
CQ		
CR		
CS		
CT		
CU		
CV		
CW		
CX		
CY		
CA		
CB		
CC		
CD		
CE		
CF		
CG		
CH		

```
QY 121 ATCTTCATTAATCTTTAGTAATAATGTTAAGTGTACAGGCGAGCTTATATAG 180
D 121 ATCTTCATTAATCTTTAGTAATAATGTTAAGTGTACAGGCGAGCTTATATAG 180
QY 181 TAAATGCGAGCTGACAGCATGAAATATACATATCTAATATTTTGTGACTATCTTATAGG 240
D 181 TAAATGCGAGCTGACAGCATGAAATATACATATCTAATATTTTGTGACTATCTTATAGG 240
QY 241 AAAATCGAGAAATTCAAAACCTTGTATGTTTAAAGGTAATAGTCACTTTTATTAATGT 300
D 241 AAAATCGAGAAATTCAAAACCTTGTATGTTTAAAGGTAATAGTCACTTTTATTAATGT 300
QY 301 GCGGTATATTTATATCATGATTTGACGTTTGTAATAATTTTCCCTGGACCTTATTTTA 360
D 301 GCGGTATATTTATATCATGATTTGACGTTTGTAATAATTTTCCCTGGACCTTATTTTA 360
QY 361 GATGAGATCTACAGTGTAGGCAAACTTATATATCTGTCACTCCATTAGTGTATAGTC 420
D 361 GATGAGATCTACAGTGTAGGCAAACTTATATATCTGTCACTCCATTAGTGTATAGTC 420
QY 421 AGACTCATCCCATGCTTAAATTTATGTTGTAATAATGCTTTTGTAAATAGTTGTGTT 480
D 421 AGACTCATCCCATGCTTAAATTTATGTTGTAATAATGCTTTTGTAAATAGTTGTGTT 480
QY 481 AGGTCAATATCAGCAAGCTTCAAGGATTAATATTAATTAACCTTGKTTTATATCTTG 540
D 481 AGGTCAATATCAGCAAGCTTCAAGGATTAATATTAATTAACCTTGKTTTATATCTTG 540
QY 541 TGAATAMCCGTTTTTTCATGCAAAATTTAAATCTTCAGCCTTTAAATTTTATTAAT 600
D 541 TGAATAMCCGTTTTTTCATGCAAAATTTAAATCTTCAGCCTTTAAATTTTATTAAT 600
QY 601 ATATAGAGATGTGATGTAGTACTACAAAACAGAAAAATTAACAGATTCGTTGTG 660
D 601 ATATAGAGATGTGATGTAGTACTACAAAACAGAAAAATTAACAGATTCGTTGTG 660
QY 661 GCTTTTGCTAAATGTTTACCGTGAACAAATCTTAGCCAGTCTTCAATTTCCCTTTGAGAT 720
D 661 GCTTTTGCTAAATGTTTACCGTGAACAAATCTTAGCCAGTCTTCAATTTCCCTTTGAGAT 720
QY 721 GAAGATATCTAGTTTATGTCAGAGGCGTGGCGGATAGTGCCTGTGTCCAGTG 780
D 721 GAAGATATCTAGTTTATGTCAGAGGCGTGGCGGATAGTGCCTGTGTCCAGTG 780
QY 781 CTTTGGCGGCGCGAGGAGGTGATCACTTAAGTCAAGAGTTTGAGACCAAGCTGCCA 840
D 781 CTTTGGCGGCGCGAGGAGGTGATCACTTAAGTCAAGAGTTTGAGACCAAGCTGCCA 840
QY 841 ACATGCGAAGCGTGTCTCTACTAATAAATTAACAAATTTAGACAGGCGTGTGGCAACA 900
D 841 ACATGCGAAGCGTGTCTCTACTAATAAATTAACAAATTTAGACAGGCGTGTGGCAACA 900
QY 901 TCTGTATATCTCAGTACTACAGAGGCTTACACAGAAAAATTCCTTGAACCTGGAGGAG 960
D 901 TCTGTATATCTCAGTACTACAGAGGCTTACACAGAAAAATTCCTTGAACCTGGAGGAG 960
QY 961 AGGTGAGTGAAGGCAATTCGACTCCAGCCTGGGCAACACAGTGAAGTCTTTGTCTCAAAA 1020
D 961 AGGTGAGTGAAGGCAATTCGACTCCAGCCTGGGCAACACAGTGAAGTCTTTGTCTCAAAA 1020
QY 1021 AAAAAAAAAAACTCGA 1038
D 1021 AAAAAAAAAAACTCGA 1038
QY 1021 AAAAAAAAAAACTCGA 1038
D 1021 AAAAAAAAAAACTCGA 1038
```

RESULT 7

```
ABD32800 standard, DNA: 68196 BP.
XX ABD32800;
XX 18-NOV-2004 (first entry)
```

```
XX Human cancer-associated genomic DNA HD16-049.
DE Human; de; cancer-associated protein; gene; cytostatic; cancer;
KW leukaemia; lymphoma; CAP.
OS Homo sapiens.
XX MO2004074320-A2.
XX 02-SEP-2004.
XX 17-FEB-2004; 2004WO-US004730.
XX 14-FEB-2003; 2003US-00367094.
XX 14-MAR-2003; 2003US-0038838.
XX 15-APR-2003; 2003US-00417375.
XX 13-JUN-2003; 2003US-00461862.
XX 15-SEP-2003; 2003US-00663431.
XX 15-DEC-2003; 2003US-00737318.
XX (SAGR-) SAGRES DISCOVERY INC.
XX Morris DW, Morris DW, Malandro MS;
XX WPI; 2004-652914/63.
XX New isolated cancer-associated polynucleotides and polypeptides useful
XX for diagnosing, preventing or treating cancer, especially lymphoma and
XX leukemia, or in screening for agents that modulate cancer.
XX claim 16; seqid 434; 310pp; English.
XX The invention relates to an isolated nucleic acid comprising at least 10
XX contiguous nucleotides of any of the 233 polynucleotide sequences given
XX in the specification, or its complement. The nucleic acids encode cancer-
XX associated proteins. Also included are an expression vector comprising
XX the isolated nucleic acid cited above, a host cell comprising the above
XX recombinant nucleic acid or expression vector, a microarray for detecting
XX a cancer-associated (CA) nucleic acid comprising at least one probe
XX comprising at least 10 contiguous nucleotides of any of the above-
XX mentioned nucleotide sequences, an isolated polypeptide (encoded within
XX an open reading frame of a CA sequence selected from any of the 95
XX polynucleotide sequences as mentioned in the specification, or its
XX complement), an isolated antibody, (or its antigen binding fragment) that
XX binds to the above polypeptide, a hybridoma that produces the above
XX monoclonal antibody, a pharmaceutical composition comprising the above
XX antibody and a pharmaceutical excipient, a kit for detecting cancer
XX cells/comprising the antibody cited above, methods for diagnosing cancer
XX or for detecting the presence or absence of cancer cells in an
XX individual, a method for inhibiting growth of cancer cells in an
XX individual, a method for delivering a therapeutic agent to cancer cells
XX in an individual, an electronic library comprising the above
XX polynucleotide or polypeptide (or their fragments), methods of screening
XX for anticancer activity or for a bioactive agent capable of modulating
XX the activity of a CA protein (CAP), methods for detecting cancer
XX associated with expression of a polypeptide in a test cell sample, a
XX method for treating cancers and a method for inhibiting the expression of
XX CA gene in a cell. The composition and methods are useful for detecting,
XX diagnosing, preventing and treating cancers, especially lymphoma and
XX leukemia. These may also be used in screening for agents that modulate
XX cancer. The present sequence is a human CAP genomic sequence. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
```

```
SQ Sequence 68196 BP; 20295 A; 14326 C; 13150 G; 20385 T; 0 U; 40 Other;
Query Match 19.3%; Score 200; DB 13; Length 68196;
Best Local Similarity 81.7%; Pred. No. 4.9e-33;
Matches 245; Conservative 0; Mismatches 50; Indels 5; Gaps 1;
QY 732 GTTTAGTCAAGGGGCTGGCGGATGATGAGCCTGTGTCTCCAGTGTCTTGGCGGGC 791
```

DB 16945 GATATTACCTTATGCTGGGACATGCTGATGCTGTATATCCCTGCATTTTGGAGGC 17004
QY 792 CGAGGACAGTGTGATCACTTAAGTTCAGAGATTGAGACCAAGCTCCCAACATGTTGAAA 851
DB 17005 CGAGGACAGTGTGATCACTTAAGTTCAGAGATTGAGACCAAGCTCCCAACATGTTGAAA 17064
QY 852 CGTTGTCTTACTTAATAAATTAATTAACAGAGCTGTGTGACACATCTGTATATCC 911
DB 17065 CTTGTCTTACTTAATAAATTAATTAACAGAGCTGTGTGACACATCTGTATATCC 17124
QY 912 AGCTACTCAGAGAGCTTAACACAGAAAATTCCTTGAACCTTGGAGGAGAGTTGCAGTG 971
DB 17125 AGCTACTCAGAGAGCTTAACAGAGAAATTCCTTGAACCTTGGAGGAGAGTTGCAGTG 17179
QY 972 AGCCATTGCACTCCAGCTGGGCAACACAGTGAAGCTTGTCTTCAAAAAATTTTAAAA 1031
DB 17180 TGGCAAAAGCACTCCAGCTGGGCAACAAATGAGACTTTGTCTCAAAAAAATTTTAAAA 17239

RESULT 8
ABT10719/c
ID ABT10719 standard; cDNA; 122748 BP.

AC ABT10719;

DT 04-DEC-2002 (first entry)

DE Human breast cancer associated coding sequence SEQ ID NO: 853.

KM Human; breast specific gene; breast cancer; differential expression;

KW cytoskeletal; gene therapy; gene; ss.

OS Homo sapiens.

PN W0200259271-A2.

PD 01-AUG-2002.

PF 25-JAN-2002; 2002W0-US0002176.

PR 25-JAN-2001; 2001US-0263757P.

PR 25-APR-2001; 2001US-0286090P.

PR 23-MAY-2001; 2001US-0292517P.

PA (GENE-) GENE LOGIC INC.

PI Orr MS, Nalion M, Diggins JC, Zeng W;

DR WPI; 2002-674803/72.

PT Diagnosing breast cancer in a patient comprises detecting the level of

PT gene expression in cell or tissue samples, where a differential gene

PS expression is indicative of breast cancer.

PS Claim 1; SEQ ID NO 853; 260Cp + Sequence Listing; English.

CC The present invention relates to methods of diagnosing breast cancer in a

CC patient, which comprise detecting the level of expression in a tissue

CC sample of two or more genes selected from those shown in ABT09867-

CC ABT1112, where a differential expression of the genes indicates breast

CC cancer. The methods are useful in diagnosing, treating, detecting the

CC progression, and in monitoring treatment of breast cancer in patients.

CC The methods are also useful as a screening tool for agents that modulate

CC the onset or progression of breast cancer. The breast cancer genes may be

CC used as diagnostic markers for the prediction or identification of the

CC malignant state of breast tissue, for confirming the type and progression

CC of cancer, and for drug screening and assays. The present sequence is a

CC coding sequence of the invention. Note: The sequence data for this patent

CC did not form part of the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub.published_pct_sequences

SEQ Sequence 122748 BP; 32088 A; 31056 C; 30547 G; 29057 T; 0 U; 0 Other;

Query Match 19.2%; Score 199.8; DB 6; Length 122748;

Best Local Similarity 82.3%; Pred. No. 6.1e-33; Index 11; Gaps 1;

Matches 246; Conservative 0; Mismatches 42;

QY 745 GGCTGGGCGCGATATGATGCTGTGTGTCCTGATGCTTGTGCGGAGCCGAGGAGGTGA 804

DB 119271 GGGTGGGTGAGTGGCTGACGCTGTATATCCAGCACTTTGGAGGTGAGGAGGTGA 119212

QY 805 TCATTTAGTCAAGAGTTTGAACACAGCTGCTCCCAACATGTTGAAAACCTGTCTACT 864

DB 119211 TCATTTAGTCAAGAGTTTGAACACAGCTGCTCCCAACATGTTGAAAACCTGTCTACT 119152

QY 865 AAAAAATCAAAAATTTGACAGAGGCTGTGACACACATCTGTATATTCAGTACTCAGAG 924

DB 119151 AAAAAATCAAAAATTTGACAGAGGCTGTGACACACATCTGTATATTCAGTACTCAGAG 119092

QY 925 GCTAACACAGAAAATTTCTTGAACCTTGGAGGACAGAGTTGCAGTGAG----- 973

DB 119091 GCTAACACAGAGAAATCTTGAACCTTGGAGGACAGAGTTGCAGTGAGCCAAAGATTGCA 119032

QY 974 CCATTGCACTCCAGCTGGGCAACACAGTGAAGCTTGTCTTCAAAAAAATTTTAAAA 1032

DB 119031 CCATTGCACTCCAGCTGGGCAACAGAGTGAAGCTTGTCTTCAAAAAAATTTTAAAA 118973

RESULT 9

ADP45591
ID ADP45591 standard; DNA; 92500 BP.

AC ADP45591;

DT 26-AUG-2004 (first entry)

DE Human intercellular adhesion molecule ICAM-1/ICAM-4/ICAM-5 gDNA.

KM breast cancer; cytoskeletal; gene therapy; human;

KW intercellular adhesion molecule; ICAM-1; human rhinovirus receptor; BB2;

KW CD54; cell surface glycoprotein P3.58; ICAM-4;

KW Landsreiner-Wiener blood group; ICAM-5; telencephalin; chromosome 19p13;

OS ds; gene; SNP; single nucleotide polymorphism.

OS Homo sapiens.

OS Key

FT variation

FT variation

FT variation

FT variation

FT variation

FT variation

FT variation

FT variation

FT variation

FT variation

FT variation

FT variation

Location/Qualifiers

39
/*tag= a

/standard_name= "Single nucleotide polymorphism (SNP) "

11799
/*tag= b

/standard_name= "Single nucleotide polymorphism (SNP) "

11851
/*tag= c

/standard_name= "Single nucleotide polymorphism (SNP) "

11963
/*tag= d

/standard_name= "Single nucleotide polymorphism (SNP) "

24282
/*tag= e

/standard_name= "Single nucleotide polymorphism (SNP) "

26849
/*tag= f

/standard_name= "Single nucleotide polymorphism (SNP) "

29633
/*tag= g

/standard_name= "Single nucleotide polymorphism (SNP) "

31254
/*tag= h

/standard_name= "Single nucleotide polymorphism (SNP) "

31967
/*tag= i

/standard_name= "Single nucleotide polymorphism (SNP) "

PT	variation	32920	/*tag= j	PT	/*tag= ah
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	45536
PT	variation	33929	/*tag= k	PT	/*tag= ai
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	46153
PT	variation	35599	/*tag= l	PT	/*tag= aj
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	47546
PT	variation	36101	/*tag= m	PT	/*tag= ak
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	47697
PT	variation	36340	/*tag= n	PT	/*tag= al
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	47944
PT	variation	36405	/*tag= o	PT	/*tag= am
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	48530
PT	variation	36517	/*tag= p	PT	/*tag= an
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	51102
PT	variation	36777	/*tag= q	PT	/*tag= ao
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	57090
PT	variation	36992	/*tag= r	PT	/*tag= ap
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	60093
PT	variation	37645	/*tag= s	PT	/*tag= aq
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	60439
PT	variation	37868	/*tag= t	PT	/*tag= ar
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	62694
PT	variation	38440	/*tag= u	PT	/*tag= as
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	66260
PT	variation	38532	/*tag= v	PT	/*tag= at
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	67295
PT	variation	38547	/*tag= w	PT	/*tag= au
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	67304
PT	variation	38712	/*tag= x	PT	/*tag= av
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	67731
PT	variation	40684	/*tag= y	PT	/*tag= aw
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	68555
PT	variation	40860	/*tag= z	PT	/*tag= ax
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	70429
PT	variation	41213	/*tag= aa	PT	/*tag= ay
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	70875
PT	variation	41419	/*tag= ab	PT	/*tag= az
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	72360
PT	variation	41613	/*tag= ac	PT	/*tag= ba
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	74228
PT	variation	42407	/*tag= ad	PT	/*tag= bb
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	76802
PT	variation	43440	/*tag= ae	PT	/*tag= bc
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	77664
PT	variation	44247	/*tag= af	PT	/*tag= bd
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	78803
PT	variation	44677	/*tag= ag	PT	/*tag= be
PT	/standard_name= "Single nucleotide polymorphism (SNP) "		/standard_name= "Single nucleotide polymorphism (SNP) "	PT	79263
PT	variation	45256	/*tag= bf	PT	

```
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT 80810
FT /*tag= bg
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT 81020
FT /*tag= bh
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT 82426
FT /*tag= bi
FT /standard_name= "Single nucleotide polymorphism (SNP)"
```

```
Query Match 19.2%; Score 199.2; DB 12; Length 92500;
Best Local Similarity 80.8%; Pred. No. 7.8e-33;
Matches 231; Conservative 1; Mismatches 54; Indels 0; Gaps 0;
```

```
QY 745 GGCTGGGCGGAGAGTGAATGCTGGTGGCCAGTCTTGGCGGCGGAGGAGTGA 804
DB 82767 GGCTGGGCGGAGTGGCTGACGCTTAATATCCAGCATTTTGGGAGGCGGAGTGA 82826
QY 805 TCACCTTAAGTCAAGAGTTTGAGACCAAGCTGCCCAACATGGTGAACGTTGCTCTACT 864
DB 82827 TCACCTGAAGTCAAGAGTTTGAGACCAAGCTGCCCAACATGGTGAACCTTGTCTTACT 82886
QY 865 AAAAAATCAAAAAATTGAGACGCGTGTGGCAACAATCTGTAATTCAGCTACTCAGAG 924
DB 82887 AAAAAATCAAAAAATTGAGACGCGTGTGGCGACACCTGTATCTTACTGAGGAG 82946
QY 925 GCTAACACAGAAAAATTCCTTGAACCTGGAGGAGGAGGTTGAGTGCATTCGACTC 984
DB 82947 GCTGAGGTGAGAGAACTGCTTCCACCCAGGAGGAGGTTGAGTGCATTCGACTC 83006
QY 985 CAGCTGGGCAACAACAGTGAAGCTTGTCTCAAAAAA 1030
DB 83007 CAGCTGGGTGACAAAGGAGGAGCACTGTCAACAATTAACA 83052
```

RESULT 10

```
ACF42745.0
WP Sequence split into 4 fragments LOCUS ACF42745 Accession ACF42745
WP Fragment Name Begin End
WP ACF42745.0 1 110000
WP ACF42745-1 100001 210000
WP ACF42745-2 200001 310000
WP ACF42745-3 300001 354391
ID ACF42745 Standard; DNA; 354391 Bp.
```

ACF42745;

29-SEP-2003 (first entry)

Human ALMS1 genomic DNA sequence.

Human; ALMS1; chromosome 2; 2p13; Alstrom disease; retinal dystrophy;
cardiomyopathy; endocrinopathy; diabetes; Alstrom syndrome; cardiant;
ophthalmological; antidiabetic; hepatotropic; nephrotropic; gene therapy;
gene; ds.

Homo sapiens.

MO2003034072-A2.

24-APR-2003.

15-OCT-2002; 2002MO-GB004658.

15-OCT-2001; 2001GB-00024621.

22-OCT-2001; 2001GB-00025318.

07-JAN-2002; 2002GB-00000248.

08-FEB-2002; 2002GB-00003039.

08-FEB-2002; 2002GB-00003040.

(UYSO-) UNIV SOUTHAMPTON.

PI Wilson DI, Hearn T, Walker M;
DR WPI; 2003-393556/37.
XX
XX
PT Diagnosing the presence of, or susceptibility to, retinal dystrophy,
PT cardiomyopathy, endocrinopathy, diabetes, or Alstrom syndrome in an
PT individual, comprises detection or modulation of the ALMS1 protein or
PT gene region.

Disclosure; Page 42-101; 121pp; English.

The present invention describes a method of diagnosing the presence of,
or susceptibility to, retinal dystrophy, cardiomyopathy, endocrinopathy,
diabetes, or Alstrom syndrome in an individual. The method comprises
typing in a sample from the individual the ALMS1 protein or ALMS1 gene
region of the individual, or detecting aberrant ALMS1 activity. Human
ALMS1 is located to chromosome 2, more specifically to 2p13. ALMS1 has
ophthalmological, cardiant, antidiabetic, hepatotropic and nephrotropic
activities, and can be used in gene therapy. The method is useful for
diagnosing the presence of, or susceptibility to, retinal dystrophy,
cardiomyopathy, endocrinopathy (e.g. liver disease or renal impairment),
diabetes, or Alstrom syndrome in an individual. ALMS1 sequences can be
used in an agent that prevents or treats retinal dystrophy,
cardiomyopathy, endocrinopathy or diabetes is useful in manufacturing a
medicament for treating a patient who has been diagnosed as having or
being susceptible to retinal dystrophy, cardiomyopathy, endocrinopathy or
diabetes. ACF42632 to ACF42747 and ABR82113 to ABR82118 represent
sequences used in the exemplification of the present invention

Sequence 354391 Bp; 98637 A; 69688 C; 70813 G; 115253 T; 0 U; 0 Other;

```
Query Match 19.2%; Score 199.2; DB 10; Length 110000;
Best Local Similarity 80.1%; Pred. No. 8.1e-33;
Matches 234; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
```

```
QY 743 GGCGCTGGGCGGAGTGAATGCTGGTGGCCAGTCTTGGCGGCGGAGGAGTGA 802
DB 28615 GGCGCGGAGGCGGTGCTTCAAGCTGTATCCAGCATTTGGGAGGCTGAGGAGGTG 28674
QY 803 GATCACTTAAGTCAAGAGTTTGAGACCAAGCTGCCCAACATGGTGAACGTTGCTCTA 862
DB 28675 GATTAACCTGAAGTCAAGAGTTCAAGACCAAGCTGGCTTACATGGTGAACCTGCTCCA 28734
QY 863 CTAAAAATCAAAAAATTGAGACGCGTGTGGCAACATCTGTAATTCAGCTACTCAGG 922
DB 28735 CTAAAAATCAAAAAATTGAGACGCGGTGTGTGGCACCGGCTTAATCCAGCTACTTGGG 28794
QY 923 AGGCTTAACAACAGAAATTCCTTGAACCTGGAGGCGAGAGTTGAGTGAAGCCATTGCA 982
DB 28795 AGGCTGAGGCGAGAGAACTCTTGAATCTGGAGGCGGAGGTGGCAGTGAAGCCATGCA 28854
QY 983 TCACGCTGGGCAACAACAGTGAAGCTTGTCTCAAAAAA 1034
DB 28855 TCCTGCTGGGTGACAGAGTGAAGCTTGTCTCAAAAAA 28906
```

RESULT 11

```
ABD32548/c
ID ABD32548 standard; DNA; 310268 Bp.
```

ABD32548;

18-NOV-2004 (first entry)

Human cancer-associated genomic DNA HD14-033.

Human; ds; cancer-associated protein; gene; cytostatic; cancer;
leukaemia; lymphoma; CAP.

Homo sapiens.

MO2004074320-A2.

PD		02-SEP-2004.
XX		
PP		17-FEB-2004; 2004WO-US004730.
XX		
PR	14-FEB-2003;	2003US-00367094.
PR	14-MAR-2003;	2003US-00388838.
PR	15-APR-2003;	2003US-00417375.
PR	13-JUN-2003;	2003US-00461862.
PR	15-SEP-2003;	2003US-00663431.
PR	15-DEC-2003;	2003US-00737318.
PA	(SAGR-) SAGRES DISCOVERY INC.	
PI	Morris DW, Morris DW, Malandro MS;	
XX	WPI; 2004-652914/63.	
DR		
XX		
PT	New isolated cancer-associated polynucleotides and polypeptides useful	
PT	for diagnosing, preventing or treating cancers, especially lymphoma and	
PT	leukemia, or in screening for agents that modulate cancer.	
XX		
PS	claim 16; seqid 24; 310bp; English.	
XX		
CC	The invention relates to an isolated nucleic acid comprising at least 10	
CC	contiguous nucleotides of any of the 233 polynucleotide sequences given	
CC	in the specification, or its complement. The nucleic acids encode cancer-	
CC	associated proteins. Also included are an expression vector comprising	
CC	the isolated nucleic acid cited above, a host cell comprising the above	
CC	recombinant nucleic acid or expression vector, a microarray for detecting	
CC	a cancer-associated (CA) nucleic acid comprising at least one probe	
CC	comprising at least 10 contiguous nucleotides of any of the above-	
CC	mentioned nucleotide sequences, an isolated polypeptide (encoded within	
CC	an open reading frame of a CA sequence selected from any of the 95	
CC	polynucleotide sequences as mentioned in the specification, or its	
CC	complement), an isolated antibody, (or its antigen binding fragment) that	
CC	binds to the above polypeptide, a hybridoma that produces the above	
CC	monoclonal antibody, a pharmaceutical composition comprising the above	
CC	antibody and a pharmaceutical excipient, a kit for detecting cancer	
CC	cells/comprising the antibody cited above, methods for diagnosing cancer	
CC	or for detecting the presence or absence of cancer cells in an	
CC	individual, a method for inhibiting growth of cancer cells in an	
CC	individual, a method for delivering a therapeutic agent to cancer cells	
CC	in an individual, an electronic library comprising the above	
CC	polynucleotide or polypeptide (or their fragments), methods of screening	
CC	for anticancer activity or for a bioactive agent capable of modulating	
CC	the activity of a CA protein (CAP), methods for detecting cancer	
CC	associated with expression of a polypeptide in a test cell sample, a	
CC	method for treating cancers and a method for inhibiting the expression of	
CC	CA gene in a cell. The composition and methods are useful for detecting,	
CC	diagnosing, preventing and treating cancers, especially lymphoma and	
CC	leukaemia. These may also be used in screening for agents that modulate	
CC	cancer. The present sequence is a human CAP genomic sequence. Note: The	
CC	sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic format directly from WIPO	
CC	at ftp.wipo.int/pub/published_pct_sequences	
XX		
SQ	Sequence 310268 BP; 87532 A; 60932 C; 62901 G; 98913 T; 0 U; 0 Other;	
	Query Match 19.2%; Score 199.2; DB 13; Length 310268;	
	Boot Local Similarity 82.0%; Pred. No. 1e-32;	
	Matches 246; Conservative 0; Mismatches 43; Indels 11; Gaps 1;	
OY	745 GGCTGGGCGGCATGTACTGTCCTGTGTCGCCAGTGTGGCGGGCCGAGCAGGTGA 804	
Dd	140417 GGCTGGCATGTTACCTTAATGCTGTAAATCCAGCACCTTTGGAGAGCCGACGCAATTGA 140418	
OY	805 TCACCTTAAGTCAGAGATTGTGAGACCAGCCTTGCCCAACATGTGAAACGTTGCTCTACT 864	
Dd	140417 TCACCTTGAAGTCAGGAATTCGAGACCAAGCCTGACCAACATGTGAAACCTGCTCTACT 140358	
OY	865 AAAAATAACAAAATTNAGACAGCGCTGTGTGGCACACATCTGTATTTCCAGTACTACGAGAG 924	
Dd	140357 AAAAATAACAAAATTNAGCAGGCAATGTGTGTGGCACCTGTATTTCCAGGACTGTGGGAG 140298	

Oy	925	CTTAACACAGGAAATTCCTTGAACCTGGGAGGACAGAGTTGACAGAG-----	973
Db	140297	GCTAAGGAGGAGAAATCCGCTTGACCTGGAGGACAGAGTTGACAGCCCAATCACA	140238
Oy	974	CCATTGCACTCCAGCTGGGCAACAAGTGAAGCTTGTCTCMAAAAAAAAAAAAAA	1033
Db	140237	CCACTGCACCTCAGCGCTGGGCAACAGATGAGACTTCGCTCCATTAAAAAAAAAAAA	140178
RESULT 12			
AA	AA	AA	AA
AL04931	standard; DNA; 32249 BP.		
XX			
AC	AA		
XX			
DT	21-NOV-2001 (first entry)		
XX			
DE	Human reproductive system related antigen DNA SEQ ID NO: 7619.		
XX			
KW	Human; reproductive system related antigen; reproductive system disorder;		
XX	cancer; gene therapy; db.		
XX			
OS	Homo sapiens.		
XX			
PN	WO20015320-A2.		
XX			
PD	02-AUG-2001.		
XX			
PF	17-JAN-2001; 2001WO-US001339.		
XX			
PR	31-JAN-2000; 2000US-0179065P.		
PR	04-FEB-2000; 2000US-0180628P.		
PR	24-FEB-2000; 2000US-0184664P.		
PR	02-MAR-2000; 2000US-0186350P.		
PR	16-MAR-2000; 2000US-0189874P.		
PR	17-MAR-2000; 2000US-0190076P.		
PR	18-APR-2000; 2000US-0198123P.		
PR	19-MAY-2000; 2000US-0205515P.		
PR	07-JUN-2000; 2000US-0209467P.		
PR	28-JUN-2000; 2000US-0214886P.		
PR	30-JUN-2000; 2000US-0215135P.		
PR	07-JUL-2000; 2000US-0216647P.		
PR	07-JUL-2000; 2000US-0216880P.		
PR	11-JUL-2000; 2000US-0217487P.		
PR	11-JUL-2000; 2000US-0217496P.		
PR	14-JUL-2000; 2000US-0218290P.		
PR	26-JUL-2000; 2000US-0220964P.		
PR	14-AUG-2000; 2000US-0224518P.		
PR	14-AUG-2000; 2000US-0224519P.		
PR	14-AUG-2000; 2000US-0225213P.		
PR	14-AUG-2000; 2000US-0225214P.		
PR	14-AUG-2000; 2000US-0225266P.		
PR	14-AUG-2000; 2000US-0225267P.		
PR	14-AUG-2000; 2000US-0225268P.		
PR	14-AUG-2000; 2000US-0225270P.		
PR	14-AUG-2000; 2000US-0225447P.		
PR	14-AUG-2000; 2000US-0225757P.		
PR	14-AUG-2000; 2000US-0225758P.		
PR	14-AUG-2000; 2000US-0225759P.		
PR	18-AUG-2000; 2000US-0226279P.		
PR	22-AUG-2000; 2000US-0226681P.		
PR	22-AUG-2000; 2000US-0226686P.		
PR	22-AUG-2000; 2000US-0227182P.		
PR	23-AUG-2000; 2000US-0227109P.		
PR	30-AUG-2000; 2000US-0228924P.		
PR	01-SEP-2000; 2000US-0229287P.		
PR	01-SEP-2000; 2000US-0229343P.		
PR	01-SEP-2000; 2000US-0229344P.		
PR	01-SEP-2000; 2000US-0229345P.		
PR	05-SEP-2000; 2000US-0229509P.		
PR	05-SEP-2000; 2000US-0229511P.		

XX Human testicular antigen encoding DNA fragment SEQ ID NO: 2477.
DE
XX
KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW reproductive system disorder; urinary system disorder; gene therapy;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disease; infection; cytostatic; gene; ds.
XX
OS Homo sapiens.
XX
PN W0200155317-A2.
XX
PD 02-AUG-2001.
XX
PE 17-JAN-2001; 2001WO-US001329.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 26-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225246P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232387P.
PR 14-SEP-2000; 2000US-0232388P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234937P.
PR 25-SEP-2000; 2000US-0234948P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249247P.
PR 17-NOV-2000; 2000US-0249255P.
PR 17-NOV-2000; 2000US-0249257P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.

```
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0254678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483232/52.
XX
XX Nucleic acids encoding 973 human testicular antigen polypeptides, useful
XX PT for preventing, diagnosing and/or treating testicular cancer.
XX
XX PS Disclosure; SEQ ID NO 2477; 766bp; English.
XX
XX CC The present invention provides the protein and coding sequences of 973
XX CC human testicular antigens, and fragments of their genomic sequences. The
XX CC sequences can be used in the treatment of cardiovascular, urinary system,
XX CC reproductive system, immune, respiratory, neurological and
XX CC gastrointestinal disorders, infections, and particularly cancer,
XX CC especially testicular cancers. The present sequence is a DNA encoding a
XX CC protein fragment of the invention
XX
XX SQ Sequence 32249 BP; 9657 A; 6826 C; 7366 G; 8400 T; 0 U; 0 Other;
XX
XX Query Match 19.1%; Score 198.2; DB 4; Length 32249;
XX Best Local Similarity 81.9%; Pred. No. 1e-32;
XX Matches 245; Conservative 0; Mismatches 43; Indels 11; Gaps 1;
XX
XX QY 745 GCGTGGCGCGGATGCTGATGCTGTGTGTCCTCCAGTCTTTGGGGCGGAGCGAGTGA 804
XX DB 11877 GGGCTGGCGCGGATGCTGATGCTGTGTGTCCTCCAGTCTTTGGGGCGGAGCGAGTGA 11936
XX
XX QY 805 TCACCTTAAGGTGAGAGTTTGAGACGAGCCCTGCAACATGTAAGAGTTGCTCTACT 864
XX DB 11937 TCACCTTAAGGTGAGAGTTTGAGACGAGCCCTGCAACATGTAAGAGTTGCTCTACT 11996
XX
XX QY 865 AAAAATCAAAAATTAGACAGGCGTGTGTGACACATCTGTATTTCCAGTACTCAGGAG 924
XX DB 11997 AAAAATCAAAAATTAGACAGGCGTGTGTGACACATCTGTATTTCCAGTACTCAGGAG 12056
XX
XX QY 925 GCTTAACACAGAAAATTCCTTGAACCTGGAGGAGAGGTTGCTGAGGCAT----- 977
XX DB 12057 GCTGAACACAGAAAATTCCTTGAACCTGGAGGAGAGGTTGCTGAGGCATGCTGCA 12116
XX
XX QY 978 -----TGACCTCCAGCTGGGCAACACAGTGAAGCTTTGTCTCAAAAAA 1032
XX DB 12117 CCACTGACCTCCAGCTGGGCAACAGTGAAGCTTTGTCTCAAAAAA 12175
XX
XX RESULT 14
XX AAL61140
XX ID AAL61140 standard; DNA; 4165 BP.
XX
XX AC AAL61140;
XX
XX DT 22-SEP-2003 (first entry)
XX
XX XX Human 76P antisense transcript DNA #3.
XX
XX DE Transcription; therapy; systemic lupus erythematosus; inflammatory disease;
XX KW cardiovascular; gastrointestinal disease; atherosclerosis; skin disease;
XX KM rheumatoid disorder; autoimmune disease; diabetes; multiple sclerosis;
XX KW cancer; neuroprotective; human; 76P; antisense; gene; ds.
XX
XX OS Homo sapiens.
XX
XX PN WO2003046220-A1.
XX
XX XX 05-JUN-2003.
```

```
XX
XX PF 11-NOV-2002; 2002MO-IL000904.
XX
XX XX 26-NOV-2001; 2001US-00993398.
XX
XX PR 24-JUL-2002; 2002US-00201605.
XX
XX XX (COMP-) COMPUGEN LTD.
XX
XX PA Levanon E, Pollock S, Nemzer S, Shoshan A, Khosravi R, Malach S;
XX PI Levine Z, Bernstein J, Dahari D, Wasserman A, Rotman G;
XX
XX DR WPI; 2003-505211/47.
XX
XX XX
XX PT Identifying putative naturally occurring antisense transcripts, useful
XX PT for quantifying gene expression levels, and detecting and/or treating
XX PT inflammatory, autoimmune or cardiovascular diseases, diabetes, arthritis
XX PT and cancer.
XX
XX PS Example 1; Page 225-227; 230pp; English.
XX
XX CC The invention relates to methods and systems for identifying naturally
XX CC occurring antisense transcripts. The method is useful for quantifying
XX CC gene expression levels and for detecting, quantifying or specifically
XX CC regulating antisense and respective sense transcripts thereby enabling
XX CC detection and treatment of a wide range or disorders, such as systemic
XX CC lupus erythematosus, inflammatory diseases, rheumatoid disorders,
XX CC autoimmune diseases, diabetes, multiple sclerosis, cardiovascular and
XX CC gastrointestinal diseases, atherosclerosis, skin diseases and cancer. The
XX CC present sequence is human 76P antisense transcript DNA used to illustrate
XX CC the method of the invention
XX
XX SQ Sequence 4165 BP; 1123 A; 994 C; 954 G; 1094 T; 0 U; 0 Other;
XX
XX Query Match 18.9%; Score 196.6; DB 9; Length 4165;
XX Best Local Similarity 66.2%; Pred. No. 1.5e-32;
XX Matches 304; Conservative 0; Mismatches 144; Indels 11; Gaps 1;
XX
XX QY 586 AATTTTATTAATTAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 645
XX DB 2672 AAGTAATTAATTAATTAATTAAGATGATGATGATGATGATGATGATGATGATGATGAT 2731
XX
XX QY 646 CAGATTCGTTTGTGCTTTGCTTAATTTGTAACCTGACAAAATCTTAGCCAGTCTTCA 705
XX DB 2732 CAGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2791
XX
XX QY 706 TTTTCTGTTGAATGAAGTACTTGAATTTTCAAGGGGCTGGGGGATGATGATGATGATGATG 765
XX DB 2792 AATTTCTAGCTAATGAATTAATTAATTAATTTTCTAGGGGCTGGGGGATGATGATGATGAT 2851
XX
XX QY 766 CCGTGGTCCAGATGCTTTGCGGGCCGAGGAGTGGATCACTTAAGTCCAGAGTTTG 825
XX DB 2852 CCGTGAATCCAGACACTTTTGGAGGCGGAGCGGGCGGATCACTTAAGTCCAGAGTTTG 2911
XX
XX QY 826 AGACCAAGCTGCCCAACATGATGTAACGTTGCTCTAATAAATAATCAAAAATTTAGCAG 885
XX DB 2912 AGACCAAGCTGCCCAACATGATGAATGTAATTTCTAATAAATAATCAAAAATTTAGCAG 2971
XX
XX QY 886 GCGTGTGGGACACATGTAATTTCCAGCTAATCCAGAGGCTAACAAGAAAATTCCTT 945
XX DB 2972 GGTGTGTGGGATGCTGTAATTTCCAGCTAATCCAGAGGCTAACAAGAAAATTCCTT 3031
XX
XX QY 946 GAACCTGGAGGAGGAGGTTGTCAGTGA-----GCCATTGCACTCCAGCTGGGC 994
XX DB 3032 GAACCTGGAGGAGGAGGTTGTCAGTGA-----GCCATTGCACTCCAGCTGGGC 3091
XX
XX QY 995 AACCAAGTGAAGTCTTTGCTCAAAAAA 1033
XX DB 3092 AACCAAGTGAAGTCTTTGCTCAAAAAA 3130
XX
XX RESULT 15
XX ADM98105
XX ID ADM98105 standard; cDNA; 4165 BP.
```

XX	AC	ADM98105;
XX	DT	17-JUN-2004 (first entry)
XX	DE	Human natural antisense sequence for 53BP1, 76p cDNA #3.
XX	KW	Human, ss; natural antisense sequence; diabetes; autoimmune disease;
XX	KW	Parkinson's disease; Alzheimer's disease; HIV infection; malaria;
XX	KW	Cholera; influenza; rabies; diphtheria; breast cancer; colon cancer;
XX	KW	cervical cancer; melanoma; lung cancer; ovarian cancer;
XX	KW	pancreatic cancer; prostate cancer; lymphoma; leukaemia.
XX	OS	Homo sapiens.
XX	PN	US2003228618-A1.
XX	PD	11-DEC-2003.
XX	PF	20-MAY-2003; 2003US-00441281.
XX	PR	24-NOV-2000; 2000US-00718407.
XX	PR	11-DEC-2000; 2000US-00732938.
XX	PR	20-FEB-2001; 2001US-00785493.
XX	PR	18-JUL-2001; 2001US-00907922.
XX	PR	26-NOV-2001; 2001US-00993329.
XX	PR	24-JUL-2002; 2002US-00201605.
XX	PR	11-NOV-2002; 2002WO-11000904.
XX	PA	(LEVA/) LEVANON E.
XX	PA	(POL/) POLLOCK S.
XX	PA	(NEMZ/) NEMZER S.
XX	PA	(SHOS/) SHOSHANI A.
XX	PA	(KHOS/) KHOSRAVI R.
XX	PA	(WALA/) WALACH S.
XX	PA	(LEVI/) LEVINE Z.
XX	PA	(BERN/) BERNSTEIN J.
XX	PA	(DAHA/) DAHARI D.
XX	PA	(WASS/) WASSERMAN A.
XX	PA	(ROTM/) ROTMAN G.
XX	PI	Levanon E, Pollock S, Nemzer S, Shoshani A, Khosravi R, Walach S;
XX	PI	Levine Z, Bernstein J, Dahari D, Wasserman A, Rotman G;
XX	PT	WPI; 2004-042222/04.
XX	PT	identifying putative natural antisense transcripts by aligning database
XX	PT	having sense-oriented polynucleotide sequence with database of expressed
XX	PT	polynucleotide sequence, identifying expressed polynucleotide sequences.
XX	PS	Example 1; SEQ ID NO 38; 133p; English.
XX	XX	
CC	CC	The invention relates to identifying putative naturally occurring
CC	CC	antisense transcripts (NAS) by computationally aligning a first database
CC	CC	of sense-oriented polynucleotide sequences with a second database of
CC	CC	expressed polynucleotide sequences, identifying expressed polynucleotides
CC	CC	forming a duplex with sense-oriented polynucleotides, or screening a
CC	CC	database of expressed polynucleotide sequences based on a sequence
CC	CC	criterion selected to identify NAS. Also included a kit for quantifying
CC	CC	an mRNA transcript of interest, a kit for quantifying an NAS of
CC	CC	interest, designing artificial antisense transcripts using a database of
CC	CC	NAS, a computer readable storage medium comprising a database of NAS, a
CC	CC	system for generating a database of several of NAS, quantifying an mRNA
CC	CC	of interest in a biological sample, quantifying the expression potential
CC	CC	of an mRNA of interest in a biological sample, quantifying NAS of
CC	CC	interest in a biological sample, identifying a novel drug target
CC	CC	(involves determining expression level of NAS in cells having an abnormal
CC	CC	phenotype and comparing the expression level of NAS of interest in cells
CC	CC	having an abnormal phenotype with an expression level cells having a
CC	CC	normal phenotype), treating or preventing a disease, condition or
CC	CC	syndrome associated with an up regulation of NAS complementary to a
CC	CC	naturally occurring mRNA transcript and diagnosing a disease, condition
CC	CC	or syndrome associated with a standard expression ratio of an mRNA of

CC interest over NAS complementary to the mRNA of interest. The method is
CC useful for identifying putative naturally occurring antisense transcripts
CC and for generating a database of NAS. An agent such as ribozyme, or an
CC oligonucleotide such as sense transcripts is useful for treating or
CC preventing a disease, condition or syndrome associated with up regulation
CC of NAS. NAS is useful for diagnosing and treating a variety of diseases
CC or pathological conditions associated with an abnormal expression of an
CC mRNA molecule of interest, comprising diabetes, autoimmune diseases,
CC Parkinson's disease, Alzheimer's disease, HIV, malaria, cholera,
CC influenza, rabies, diphtheria, breast cancer, colon cancer, cervical
CC cancer, melanoma, lung cancer, ovarian cancer, pancreatic cancer,
CC prostate cancer, lymphoma and leukemia. The present sequence is a human
CC natural antisense transcript identified by the method of the invention.

Query Match	18.9%	Score 196.6;	DB 12;	Length 4165;
Best Local Similarity	66.2%	Pred. No. 1.5e-32;		
Matches 304; Conservative	0;	Mismatches 144;	Indels 11;	Gaps 1;

[illegible]

Search completed: March 6, 2005, 04:26:16
Job time : 608 secs

Db	83054	CAAAACCTTGTTAGTTTTTATAGGGTATAGTCACATTTTATTAAGTGGGATATTTATAC	83113
Oy	316	ATGATTTGACGTTTGTGMAAATATTTTTCCCTGACCTTTATTTTTAGATGAGATCTACAGT	83175
Db	83114	ATGATTTGACGTTTGTGMAAATATTTTTCCCTGACCTTTATTTTTAGATGAGATCTACAGT	83173
Oy	376	GGAGGCAACCTTATATTAATCTGCAACCTCATAGTGTCTACTGACATCATCCCATG	83175
Db	83174	GTAGGCAACCTTATATTAATCTGCAACCTCATAGTGTCTACTGACATCATCCCATG	83233
Oy	436	CTAAATATTAGTTGTGKAAATACCGCTTTGTAAATAGTGTGTAGTGCATTATACCA	83293
Db	83234	CTAAATATTAGTTGTGTTAAATACCGCTTTGTAAATAGTGTGTAGTGCATTATACCA	83293
Oy	486	AGCTTTCAAGGAKTTTCATTATTAATAAACCCTTGSKTTTTATTCTTGTGATATMCCGTTTT	83353
Db	83294	AGCTTTCAAGGATATACATTATTAATAAACCCTTGSKTTTTATTCTTGTGAAATACCGTTTT	83353
Oy	556	TCCATGCAAAAGTTAAATAATCTTCAGGCTTTAATTTTTTTATTAATAATAAAGATGTGAT	83413
Db	83354	TCCATGCAAAAGTTAAATAATCTTCAGGCTTTAATTTTTTTATTAATAATAAAGATGTGAT	83413
Oy	616	GAGTATGACTACAAAACAGAAAAAATAAACAAGTTTGGTTTGGCTTTTGCTAAATTG	83473
Db	83414	GAGTATGACTACAAAACAGAAAAAATAAACAAGTTTGGTTTGGCTTTTGCTAAATTG	83473
Oy	676	TTACCTGCAAAAATCTTAAGCAGTTCTTCAATTTGGTTTGAATGAAGTACTTTAGTT	83533
Db	83474	TTACCTGCAAAAATCTTAAGCAGTTCTTCAATTTGGTTTGAATGAAGTACTTTAGTT	83533
Oy	736	TAGTCACAGGGGCTGGGCGGATAGCTGATGCTCTGTGGTCCCAAGTGTCTTTCGGGGCCGAG	83593
Db	83534	TAGTCACAGGGGCTGGGCGGATAGCTGATGCTCTGTGGTCCCAAGTGTCTTTCGGGGCCGAG	83593
Oy	786	GCAAGTGATCTACTTAAGTCAGGAGTTTGAACAACAGCTCTGCCCAATGTGTGAACGTT	83653
Db	83594	GCAAGTGATCTACTTAAGTCAGGAGTTTGAACAACAGCTCTGCCCAATGTGTGAACGTT	83653
Oy	856	GTCTCTACTTAAATAATCAAAAAATTTAACAAGGCGTGGTGACACATCTGTAAATTCAGCT	83713
Db	83654	GTCTCTACTTAAATAATCAAAAAATTTAACAAGGCGTGGTGACACATCTGTAAATTCAGCT	83713
Oy	916	ACTCAGAGGCTTAACAACAGAAAAATTCCTTGAACCTTGGAGGACAGAGTTTGACGTGAGCC	83773
Db	83714	ACTCAGAGGCTTAACAACAGAAAAATTCCTTGAACCTTGGAGGACAGAGTTTGACGTGAGCC	83773
Oy	976	ATTGCACTTCACGCTGGGCAACAACAGTGAACCTTTGTCTCAAAAAAATTTTTAAAAAAT	83833
Db	83774	ATTGCACTTCACGCTGGGCAACAACAGTGAACCTTTGTCTCAAAAAAATTTTTAAAAAAT	83833

RESULT 2
 US-09-949-016-16554
 Sequence 16554, Application US/09949016
 Patent No. 6812339
 GENERAL INFORMATION:
 APPLICANT: VENTER, J. Craig et al.
 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 FILE REFERENCE: CLO01307
 CURRENT APPLICATION NUMBER: US/09/949,016
 CURRENT FILING DATE: 2000-04-14
 PRIOR APPLICATION NUMBER: 60/241,755
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/237,768
 PRIOR FILING DATE: 2000-10-03
 PRIOR APPLICATION NUMBER: 60/231,498
 PRIOR FILING DATE: 2000-09-08
 NUMBER OF SEQ ID NOS: 207012
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 16554
 LENGTH: 168105

```

; TYPE: DNA
; ORGANISM: Human
;
; FEATURE:
;
; NAME/KEY: misc_feature
; LOCATION: (1)..(168105)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-16554

```

Query Match	97.5%;	Score 1011.6;	DB 4;	Length 168105;
Best Local Similarity	99.18;	Pred. No. 8.5e-237;		
Matches 1011;	Conservative	5;	Mismatches 4;	Indels 0; Gaps 0;
QY	16	TTAATGATATAAATATTCTATAATGAAATTTTAATGGAATTAGAGCATATGAAAAA	75	
Db	82814	TTAATGATATAAATATTCTATAATGAAATTTTAATGGAATTAGAGCATATGAAAAA	8287	
QY	76	TGCTCTTACTGTTGAAAACATATTTGTACATTTTGGTCACTAATCTTCAATACTT	135	
Db	82874	TGCTCTTACTGTTGAAAACATATTTGTACATTTTGGTCACTAATCTTCAATACTT	8293	
QY	136	TTAAGTACTTAATAGTTAAAGTTAGACAGAGGAGCTTATATAGTAATAGGAGCTGAC	195	
Db	82934	TTAATTAACCTAATAGTTAAAGTTAGACAGAGGAGCTTATATAGTAATAGGAGCTGAC	8299	
QY	196	AGCATGAAAATTAACATATCTAATATTTTGTACTATCTTATTAAGAAAATCAGAAATTT	255	
Db	82994	AGCATGAAAATTAACATATCTAATATTTTGTAGATCTTATTAAGAAAATCAGAAATTT	8305	
QY	256	CAAAACCTGTGATGTTTTTAGGGGTATAGTCACATTTTATATAATGTGGGTATATTATAC	315	
Db	83054	CAAAACCTGTGATGTTTTTAGGGGTATAGTCACATTTTATATAATGTGGGTATATTATAC	8311	
QY	316	ATGATTTGAGGTTTGCTGMAAATATTTTCCGTGACCTTTATTTTGAATGAGATCTACGT	375	
Db	83114	ATGATTTGAGGTTTGCTGMAAATATTTTCCGTGACCTTTATTTTGAATGAGATCTACGT	8317	
QY	376	GTAGGCAAACTTATATATATCTGTCACTCCATTAGTGTCAAGTCACACTATCCCATG	435	
Db	83174	GTAGGCAAACTTATATATATCTGTCACTCCATTAGTGTCAAGTCACACTATCCCATG	8323	
QY	436	CTAAAAATTATAGTTGTAAAAATACGCTTTTGTAATAGTTGTGTAGTCAATTATACCA	495	
Db	83234	CTAAAAATTATAGTTGTAAAAATACGCTTTTGTAATAGTTGTGTAGTCAATTATACCA	8329	
QY	496	AGTCTTCAGGKATTAATATATAAATCCTTGKCTTTATCTTGTAATAMCGCTTTT	555	
Db	83294	AGTCTTCAGGKATTAATATATAAATCCTTGKCTTTATCTTGTAATAMCGCTTTT	8335	
QY	556	TCCATGCAAAATTAATTTCTTCAGCCTTTAATTTTTTATATATATATAGATGTGAT	615	
Db	83354	TCCATGCAAAATTAATTTCTTCAGCCTTTAATTTTTTATATATATATAGATGTGAT	8341	
QY	616	GAGTATGATCAAAACAGAAAAAATAAACAGATTTGCTTTGTGGCTTTTGTCTAAATTG	675	
Db	83414	GAGTATGATCAAAACAGAAAAAATAAACAGATTTGCTTTGTGGCTTTTGTCTAAATTG	8347	
QY	676	TTAATCTGAAGAAATCTTAGCCAGTCTTCAATTTTGGTTTGAAGTGAAGATCTTAAGTT	735	
Db	83474	TTAATCTGAAGAAATCTTAGCCAGTCTTCAATTTTGGTTTGAAGTGAAGATCTTAAGTT	8353	
QY	736	TAGTCCAGGGGCTGTGGCGCATAGCTGATCCCTGTGTGCCAGTCTTTGCGGGGCGAG	795	
Db	83534	TAGTCCAGGGGCTGTGGCGCATAGCTGATCCCTGTATCCCAATCTTTGCGGGGCGAG	8359	
QY	796	GCAAGTGGATCACTTAAGGTACAGAGTTTGAAGCAAGCTGCGCCACATGTGTAAACGTT	855	
Db	83594	GCAAGTGGATCACTTAAGGTACAGAGTTTGAAGCAAGCTGCGCCACATGTGTAAACGTT	8365	
QY	856	GTCCTACTAAAAATCAAAAAATTAGACAGGCGGTGGGCAACATCTGTAAATTCACGCT	915	
Db	83654	GTCCTACTAAAAATCAAAAAATTAGACAGGCGGTGGGCAACATCTGTAAATTCACGCT	8371	
QY	916	ACTCAGGAGCTTAACAAGAAAAATTCCTTGAACCTTGAGGAGCAGAGGTTGCAGTAGCC	975	

Db 83714 ACTCAGAGGCTAACACAGGAAATTCCTTGAACCTGGAGGACAGAGTTGCACTGAGCC 83773
Qy 976 ATTGCATCTCCAGCCTGGGCAACACAGTGAGACTCTTGTCTCAAAAAAAAAAAAAAACT 1035
Db 83774 ATTGCATCTCCAGCCTGGGCAACACAGTGAGACTCTTGTCTCAAAAAAAAAAAAAAAAT 83833

RESULT 3

US-09-949-016-170431/c
; Sequence 170431, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ. ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ. ID NO 170431
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-170431

Query Match 57.7%; Score 598.6; DB 4; Length 601;
Best Local Similarity 99.0%; Pred. No. 7.8e-137;
Matches 595; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 62 CATCATGAAAAAATGCTTACTGTGAAAAACATTTATTTGGTCAATTTGGTCAACATA 121
Db 601 CATCATGAAAAAATGCTTACTGTGAAAAACATTTATTTGGTCAATTTGGTCAACATA 542
Qy 122 TCTTTCAATACTTTAGTAACTAATAGTTAGTTGACAGTGAGCTTATATAGT 181
Db 541 TCTTTCAATACTTTAGTAACTAATAGTTAGTTGACAGTGAGCTTATATAGT 482
Qy 182 AATGCGAGCTGACAGCATGAAAAATCAATCTAATATTTTGTGACTATCTTAATGGA 241
Db 481 AATGCGAGCTGACAGCATGAAAAATCAATCTAATATTTTGTGACTATCTTAATGGA 422
Qy 242 AATTCAGAAATTTCAAACTGTTAGTTTGGGTATAGTACATTTATATAATGTG 301
Db 421 AATTCAGAAATTTCAAACTGTTAGTTTGGGTATAGTACATTTATATAATGTG 362
Qy 302 CGGATATTTATACATGATTTGACGTTGTGMAAATATTTTCCCTGACATTTATTTAG 361
Db 361 CGGATATTTATACATGATTTGACGTTGTGMAAATATTTTCCCTGACATTTATTTAG 302
Qy 362 ATGAGATCTACAGTGTAGGCAAACTTAATATCTGTCAACTCATTAGTGTCAATGCA 421
Db 301 RTGAGATCTACAGTGTAGGCAAACTTAATATCTGTCAACTCATTAGTGTCAATGCA 242
Qy 422 GACTCATCCCAAGCTAAATTAATAGTTGTAAATAGCGTTTGTAAATAGTTGTGTA 481
Db 241 GACTCATCCCAAGCTAAATTAATAGTTGTAAATAGCGTTTGTAAATAGTTGTGTA 182
Qy 482 GGTCAATATGACCAAGCTTCTCAAGKATACATTATATAAAACCTTGKTTTATCTTGT 541
Db 181 GGTCAATATGACCAAGCTTCTCAAGKATACATTATATAAAACCTTGKTTTATCTTGT 122
Qy 542 GAATACCGTTTTTTCATGCAAAAGTTAAATTTCTTCAAGCTTTAATTTTATTAATA 601
Db 121 GAATACCGTTTTTTCATGCAAAAGTTAAATTTCTTCAAGCTTTAATTTTATTAATA 62

Qy 602 TATAGATGTGATGACTATGACTACAAAACAGGAAAAATTAACAGATTTCTTTGTGG 661
Db 61 TATAGATGTGATGACTATGACTACAAAACAGGAAAAATTAACAGATTTCTTTGTGG 2
Qy 662 C 662
Db 1 C 1

RESULT 4

US-09-949-016-12848
; Sequence 12848, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ. ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ. ID NO 12848
; LENGTH: 51620
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(51620)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12848

Query Match 19.3%; Score 200; DB 4; Length 51620;
Best Local Similarity 81.7%; Pred. No. 7.4e-39;
Matches 245; Conservative 0; Mismatches 50; Indels 5; Gaps 1;

Qy 732 GTTTTATCCAGGGGCTGGGCGCATAGCTGATCCTGTGTGCCAGTCTTTCGGGGC 791
Db 8929 GTATTTACCTTTAGGCTGGGCGCATAGCTGATCCTGTGTGCCAGTCTTTCGGGGC 8988
Qy 792 CGAGGCGGTGATCACTTAAGTCAAGAGTTTGAACCAAGCTGCGCAACATGTTGAA 851
Db 8989 CGAGGCGGTGATCACTTAAGTCAAGAGTTTGAACCAAGCTGCGCAACATGTTGAA 9048
Qy 852 CGTTGTCTTCACTTAAATTAACAAAAATTAGACAGGCGGTGTGGCACATCTGTATTTCC 911
Db 9049 CGTTGTCTTCACTTAAATTAACAAAAATTAGACAGGCGGTGTGGGTGTGTGATTCCTC 9108
Qy 912 AGCTACTCAGAGGCTTAACACAGGAAATTCCTTGAACCTGGAGGAGGAGTTGCAAGT 971
Db 9109 AGCTACTCAGAGGCTTAACACAGGAAATTCCTTGAACCTGGAGGAGGAG-----ATTG 9153
Qy 972 AGCCATTGCACTCCAGCTGGGCAACACAGTGAGACTTGTCTCAAAAAAAAAAAAAAA 1031
Db 9164 TGCCAAAGCACTCCAGCTGGGCAACAAATGAGACTTGTGTCTCAAAAAAAAAAAAAAA 9223

RESULT 5

US-09-949-016-16503
; Sequence 16503, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016

```

; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 16503
; LENGTH: 51621
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(51621)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16503
```

```

Query Match          19.3%; Score 200; DB 4; Length 51621;
Best Local Similarity 81.7%; Pred. No. 7,4e-39;
Matches 245; Conservative 0; Mismatches 50; Indels 5; Gaps 1;
```

```

QY 732 GTTTAGTCCAGGGCTGGGCGCATGAGTGGCTGTGTCCTCCAGTCTTTGGCGGCG 791
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 8929 GTATTATCCCTTAGGCTGGGCGCATGAGTGGCTGTGTCCTCCAGTCTTTGGGAGGC 8988
QY 792 CGAGGCGAGTGGATCACTTAAGGTGAGGATTGAGACAGCCTGCCCAACATGTGTAAA 851
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 8969 CGAGGCGAGTGGATCACTTAAGGTGAGGATTGAGACAGCCTGCCCAACATGTGTAAA 9048
QY 852 CGTTGCTCTACTAATAAAATACAAAATTGACAGGCGGTGTGGCACAATCTGTAATTC 911
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9049 CTTGCTCTACTAATAAAATACAAAATTGACAGGCGGTGTGGCACAATCTGTAATTC 9108
QY 912 AGCTACTAGAGGCTTAACACAGGAAATTCCTTGAACCTGGAGGCGAGGTTGCAAGTG 971
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9109 AGCTACTAGAGGCTTAACACAGGAAATTCCTTGAACCTGGAGGCGAGGTTGCAAGTG 9163
QY 972 AGCCATTGCACTCCAGCCTGGGCAACACAGTGAAGCTTTGTCCAAAAAATTTTAAAA 1031
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9164 TGCCAAAGCACTCCAGCCTGGGCAACAAATGAGACTTTGTGTCCAAAAAATTTTAAAA 9223
```

```

RESULT 6
US-09-949-016-15231/c
; Sequence 15231, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 15231
; LENGTH: 75799
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15231
```

```

Query Match          19.3%; Score 200; DB 4; Length 75799;
Best Local Similarity 81.7%; Pred. No. 8,4e-39;
Matches 245; Conservative 0; Mismatches 50; Indels 5; Gaps 1;
```

```

QY 732 GTTTAGTCCAGGGCTGGGCGCATGAGTGGCTGTGTCCTCCAGTCTTTGGCGGCG 791
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9102 GTATTATCCCTTAGGCTGGGCGCATGAGTGGCTGTGTCCTCCAGTCTTTGGGAGGC 9043
QY 792 CGAGGCGAGTGGATCACTTAAGGTGAGGATTGAGACAGCCTGCCCAACATGTGTAAA 851
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9042 CGAGGCGAGTGGATCACTTAAGGTGAGGATTGAGACAGCCTGCCCAACATGTGTAAA 8983
QY 852 CGTTGCTCTACTAATAAAATACAAAATTGACAGGCGGTGTGGCACAATCTGTAATTC 911
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 8982 CTTGCTCTACTAATAAAATACAAAATTGACAGGCGGTGTGGCACAATCTGTAATTC 8923
QY 912 AGCTACTAGAGGCTTAACACAGGAAATTCCTTGAACCTGGAGGCGAGGTTGCAAGTG 971
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 8922 AGCTACTAGAGGCTTAACACAGGAAATTCCTTGAACCTGGAGGCGAGGTTGCAAGTG 8868
QY 972 AGCCATTGCACTCCAGCCTGGGCAACACAGTGAAGCTTTGTCCAAAAAATTTTAAAA 1031
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 8867 TGCCAAAGCACTCCAGCCTGGGCAACAAATGAGACTTTGTGTCCAAAAAATTTTAAAA 8808
```

```

RESULT 7
US-09-949-016-15449/c
; Sequence 15449, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 15449
; LENGTH: 14345
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15449
```

```

Query Match          19.2%; Score 198.8; DB 4; Length 14345;
Best Local Similarity 76.9%; Pred. No. 9,2e-39;
Matches 260; Conservative 0; Mismatches 67; Indels 11; Gaps 1;
```

```

QY 707 TTTGCTTTGAGATGATAGTACTTGTGTTGTCAGAGGGCTGGGCGGATGATGATGC 766
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9994 TTAAGTTTCTTTGTTAAGCATTTTGAATTTGAGATAGGCTGGGCGGATGATGATGC 9935
QY 767 CTGTGCTCCAGTGTCTTGGGCGGCGAGGAGTGTATCACTTAAGTCAAGAGTTTGA 826
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9934 CTGTATCCAGCACTTTGGGAGGCGGAGGTGGGTGATCACTTGAAGTCAAGAGTTTGA 9875
QY 827 GACGAGCCTGCGCCCAACATGTGTAAACCTTTGCTCTACTAATAAAATTTGACAGG 886
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9874 GACGAGCCTGCGCCCAACATGTGTAAACCTTTGCTCTACTAATAAAATTTGACAGG 9815
QY 887 CGTGTGTCACACATCTGTAATTCAGTCACTCAGAGGCTTAACAGAGAAATTCCTGTG 946
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9814 CATGTGTGCAATGACCTTGTATCTCCAGTACTTTGGAGGCTGAGGACAGAAATTCGCTTG 9755
QY 947 AACCTGAGGAGGAGGTTGAGTAG-----CCATTGCACTCCAGCCTGGGCA 995
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9754 AACCCAGAGGCGGAGGTTGCAATGAGCCAAAGATTGCAACATTCACCTCCAGCCTGGGTG 9695
QY 996 ACACAGTGAGACTCTTGTCTCAAAAAAATTTTAAAA 1033
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9694 ACAGAGCGAGACTCATCTCAAAAAAATTTTAAAA 9657
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


RESULT 8
US-09-949-016-12316
Sequence 12316, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTUR, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12316
LENGTH: 39243
TYPE: DNA
ORGANISM: Human
US-09-949-016-12316

Query Match 19.2%; Score 198.8; DB 4; Length 39243;
Best Local Similarity 76.9%; Pred. No. 1.3e-38;

Matches 260; Conservative 0; Mismatches 67; Indels 11; Gaps 1;

Qy 707 TTTCGTTTGAGATGAGATCTAGTTTATGTCACAGGCGCGCGATAGCTGATGC 766
Db 38368 TTAAGTTTCTGTTAAGCATTTTAGAATTGAGAAATAGGTGGCGCGGTCTCAGC 38427
Qy 767 CTGTGTCCAGATGCTTTGGCGGCGCGAGCGAGTGATCACTTAAGTCAGAGTTTGA 826
Db 38428 CTGTAAATCCAGACACTTTGGGAGCGCGAGGTGGTGTATCATCTGAGTCAAGATTGA 38487
Qy 827 GACCAAGCTGCCCAACATGCTGTAACCTTGTCTTACTTAAATAACAAATTTAGACAG 886
Db 38488 GACCAAGCTGCCCAACATGCTGTAACCTTGTCTTACTTAAATAACAAATTTAGCAGG 38547
Qy 887 CGTGTGCGACACATCTGTATTCACACTACTCAGAGGCTAACACAGGAAATTCCTTG 946
Db 38548 CATGTGTGCGACACCTGTATTCACACTACTCAGAGGCTAACACAGGAAATTCGCTTG 38607
Qy 947 AACCTGGAGGAGAGGTTGCACTGAG-----CCATTGCATCCAGCTGGGCA 995
Db 38608 AACCTGGAGGAGGTTGCACTGAGCAATGACCAAGATTGACCATTCACCTGGGTG 38667
Qy 996 ACACAGTGAAGCTTGTCTCAAAAAAAAAAAAAAAAAAAAA 1033
Db 38668 ACACAGTGAAGCTTGTCTCAAAAAAAAAAAAAAAAAAAAA 38705

RESULT 9
US-09-949-016-15443
Sequence 15443, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTUR, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15443
LENGTH: 39243
TYPE: DNA
ORGANISM: Human
US-09-949-016-15443

Query Match 19.2%; Score 198.8; DB 4; Length 39243;
Best Local Similarity 76.9%; Pred. No. 1.3e-38;

Matches 260; Conservative 0; Mismatches 67; Indels 11; Gaps 1;

Qy 707 TTTCGTTTGAGATGAGATCTAGTTTATGTCACAGGCGCGCGATAGCTGATGC 766
Db 38368 TTAAGTTTCTGTTAAGCATTTTAGAATTGAGAAATAGGTGGCGCGGTCTCAGC 38427
Qy 767 CTGTGTCCAGATGCTTTGGCGGCGCGAGCGAGTGATCACTTAAGTCAGAGTTTGA 826
Db 38428 CTGTAAATCCAGACACTTTGGGAGCGCGAGGTGGTGTATCATCTGAGTCAAGATTGA 38487
Qy 827 GACCAAGCTGCCCAACATGCTGTAACCTTGTCTTACTTAAATAACAAATTTAGACAG 886
Db 38488 GACCAAGCTGCCCAACATGCTGTAACCTTGTCTTACTTAAATAACAAATTTAGCAGG 38547
Qy 887 CGTGTGCGACACATCTGTATTCACACTACTCAGAGGCTAACACAGGAAATTCCTTG 946
Db 38548 CATGTGTGCGACACCTGTATTCACACTACTCAGAGGCTAACACAGGAAATTCGCTTG 38607
Qy 947 AACCTGGAGGAGAGGTTGCACTGAG-----CCATTGCATCCAGCTGGGCA 995
Db 38608 AACCTGGAGGAGGTTGCACTGAGCAATGACCAAGATTGACCATTCACCTGGGTG 38667
Qy 996 ACACAGTGAAGCTTGTCTCAAAAAAAAAAAAAAAAAAAAA 1033
Db 38668 ACACAGTGAAGCTTGTCTCAAAAAAAAAAAAAAAAAAAAA 38705

RESULT 10
US-09-949-016-131748
Sequence 131748, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTUR, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 131748
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-131748

Query Match 19.1%; Score 198.4; DB 4; Length 601;
Best Local Similarity 76.6%; Pred. No. 3.8e-39;

Matches 259; Conservative 1; Mismatches 67; Indels 11; Gaps 1;

Qy 707 TTTCGTTTGAGATGAGATCTAGTTTATGTCACAGGCGCGCGATAGCTGATGC 766
Db 256 TTAAGTTTCTGTTAAGCATTTTAGAATTGAGAAATAGGTGGCGCGGTCTCAGC 315
Qy 767 CTGTGTCCAGATGCTTTGGCGGCGCGAGCGAGTGATCACTTAAGTCAGAGTTTGA 826
Db 316 CTGTAAATCCAGACACTTTGGGAGCGCGAGGTGGTGTATCATCTGAGTCAAGATTGA 375

```
QY      827 GACACGCTGCGCAACATGTGGAACGTTGCTCTACTATAAATAATTAACAAAATTGACAGG 886
      |||
      376 GACCAAGCTGGGCAACATGTGGAACGTTGCTCTACTATAAATAATTAACAAAATTGACAGG 435
QY      887 CGTGTGTGGCAACATGTGTAATTCAGCTACTCAGAGAGGCTTAACAGAAATAATCTCTTG 946
      |||
      436 CATGTGTGTCATGACCTGTATGCTCCAGCTACTTGGAGGCTGAGGAGAAATGCTTG 495
QY      947 AACCTGGAGGAGAGAGTTTGCATGAG-----CCATTGCACTCCAGCTTGAGCA 995
      |||
      496 AACCCAGAGAGGCGAGGTTGCAATGAGCCAAAGATTGACCATGCACTCCAGCTTGAGTG 555
QY      996 ACACAGTGAAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAA 1033
      |||
      556 ACAGAGGAGAGACTCCATCTCAAAAAAAAAAAAAAAAAAAAA 593
      |||
```

```
RESULT 11
US-09-949-016-12758
; Sequence 12758, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12758
; LENGTH: 88490
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12758
```

```
Query Match      19.1%; Score 198.2; DB 4; Length 88490;
Best Local Similarity 81.9%; Pred. No. 2.4e-38;
Matches 245; Conservative 0; Mismatches 43; Indels 11; Gaps 1;

QY      745 GAGCTGGGCGCATAGCTGATGCTGTGTGCTCCAGTGGCTTGGGCGCGAGGCGAGTGA 804
      |||
      7569 GAGCTGGGCGCTGTGTGCTATGCCGTAAATCCAGCACTTGGGAGGCCAAAGCAGGTGA 7628
QY      805 TCACCTTAAGTCAAGAGTTTGAGACCAAGCCTGCCCAACATGTGAAAGTGTCTCTACT 864
      |||
      7629 TCAGTTGAGTCAAGAGCTTGAGACCAAGCCTGCCCAACATGTGAAAGTGTCTCTACT 7688
QY      865 AAAAAATCAAAAAATTGACAGGCGCTGTGTGAGACACATCTGTAATTCAGTACTCAGGAG 924
      |||
      7689 AAAAAATCAAAAAATTGACAGGAGTGTGTGAGGCGCACTTGTAAATCCAGCTACTCGGAG 7748
QY      925 GCTTAACAGAGAAATTCCTTGAACCTGGAGGAGAGGTTGCACTGA-----G 973
      |||
      7749 GCTAGGCAAGAGAAATGCTTGAACCAAGAGGTTGAGAGTTCAGTGAAGCCGAGATCATG 7808
QY      974 CCATTGCACTCCAGCTTGCGGCAACAGTGAAGCTTGTCTCAAAAAAAAAAAAAAAAAAAAA 1032
      |||
      7809 CCGTGCACTCCAGCTTGCGGTTGACAGAGCGAGACTCTTGTCTCAAAAAAAAAAAAAAAAAAAAA 7867
      |||
```

```
RESULT 12
US-09-949-016-14222
; Sequence 14222, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14222
; LENGTH: 88736
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14222
```

```
Query Match      19.1%; Score 198.2; DB 4; Length 88736;
Best Local Similarity 81.9%; Pred. No. 2.4e-38;
Matches 245; Conservative 0; Mismatches 43; Indels 11; Gaps 1;

QY      745 GAGCTGGGCGCATAGCTGATGCTGTGTGCTCCAGTGGCTTGGGCGCGAGGCGAGTGA 804
      |||
      7815 GAGCTGGGCGTGTGTGCTCATGCCGTAAATCCAGCACTTGGGAGGCCAAAGCAGGTGA 7874
QY      805 TCACCTTAAGTCAAGAGTTTGAGACCAAGCCTGCCCAACATGTGAAAGTGTCTCTACT 864
      |||
      7875 TCAGTTGAGTCAAGAGCTTGAACCAAGCCTGCCCAACATGTGAAAGTGTCTCTACT 7934
QY      865 AAAAAATCAAAAAATTGACAGGCGCTGTGTGAGACACATCTGTAATTCAGTACTCAGGAG 924
      |||
      7935 AAAAAATCAAAAAATTGACAGGAGTGTGTGCGGAGCACTGTATCCAGTACTCGGAG 7994
QY      925 GCTTAACAGAGAAATTCCTTGAACCTGGAGGAGAGGTTGCACTGA-----G 973
      |||
      7995 GCTAGGCAAGAGAAATGCTTGAACCAAGAGGTTGAGAGTTCAGTGAAGCCGAGATCATG 8054
QY      974 CCATTGCACTCCAGCTTGCGGCAACAGTGAAGCTTGTCTCAAAAAAAAAAAAAAAAAAAAA 1032
      |||
      8055 CCGTGCACTCCAGCTTGCGGTTGACAGAGCGAGACTCTTGTCTCAAAAAAAAAAAAAAAAAAAAA 8113
      |||
```

```
RESULT 13
US-09-949-016-45906/c
; Sequence 45906, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45906
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-45906
```

```
Query Match      19.0%; Score 196.8; DB 4; Length 601;
Best Local Similarity 79.8%; Pred. No. 9.4e-39;
Matches 249; Conservative 0; Mismatches 52; Indels 11; Gaps 1;
```

```
Qy 733 TTTTATGTCAGAGGCTGGGCGCATAGCTGATGCTGNGGTTCCAGTGTCTTGCGGGGCC 792
Db 596 TTATATTTTATGAGCCGGGCGATGCTGCTCAGCGCTTATCTCCAGACCTTCGGAGGCT 537
Qy 793 GAGGAGGTGATCACTTAAAGTTCAGAGTTTGAACAGCCTTCCCAACATGTTGAAC 852
Db 536 GAGGTAGCGGATCACTGAGGTTCAGAGTTTCAGACAGCTTGGCAACATGTTGAAT 477
Qy 853 GTTGTCTCTCTAAATAATTAACAGAGCGGTGGTGGACACATCTGTATTCGA 912
Db 476 CCGTCCCTACTTAAATAATTAACAGAGTATGTCGATGTCCTGTATTCGA 417
Qy 913 GCTACTGAGAGGCTTACAGAGAAATTCCTTGAACCTGGAGGAGAGGTTGCACTGA 972
Db 416 GCTACTGAGAGGCTTACAGAGAAATTCCTTGAACCTGGAGGAGAGGTTGCACTGA 357
Qy 973 GC-----CATTCGACTCCAGCCTGGCAACACAGTGAAGCTCTTGTCAAAA 1021
Db 356 GCGGAGATCGCATCACTGCACTCCAGCCTAGTGAACAGAGTGAAGACTCTTGTCAAAA 297
Qy 1022 AAAAAAAAAA 1033
Db 296 AAAAAAAAAA 285

RESULT 14
US-09-949-016-13052
; Sequence 13052, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13052
; LENGTH: 20441
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)..(20441)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13052

Query Match 19.0%; Score 196.8; DB 4; Length 20441;
Best Local Similarity 79.8%; Pred. No. 3.2e-38;
Matches 249; Conservative 0; Mismatches 52; Indels 11; Gaps 1;

Qy 733 TTTTATGTCAGAGGCTGGGCGCATAGCTGATGCTGNGGTTCCAGTGTCTTGCGGGGCC 792
Db 11684 TTATATTTTATGAGCCGGGCGATGCTGCTCAGCGCTTATCTCCAGACCTTCGGAGGCT 11743
Qy 793 GAGGAGGTGATCACTTAAAGTTCAGAGTTTGAACAGCCTTCCCAACATGTTGAAC 852
Db 11744 GAGGTAGCGGATCACTGAGGTTCAGAGTTTCAGACAGCTTGGCAACATGTTGAAT 11803
Qy 853 GTTGTCTCTCTAAATAATTAACAGAGCGGTGGTGGACACATCTGTATTCGA 912
Db 11804 CCGTCCCTACTTAAATAATTAACAGAGTATGTCGATGTCCTGTATTCGA 11863
Qy 913 GCTACTGAGAGGCTTACAGAGAAATTCCTTGAACCTGGAGGAGAGGTTGCACTGA 972
Db 913 GCTACTGAGAGGCTTACAGAGAAATTCCTTGAACCTGGAGGAGAGGTTGCACTGA 972
```

```
Db 11864 GCTACTGAGAGGCTTACAGAGAAATTCCTTGAACCTGGAGGAGAGGTTGCACTGA 11923
Qy 973 GC-----CATTCGACTCCAGCCTGGCAACACAGTGAAGCTCTTGTCAAAA 1021
Db 11924 GCGGAGATCGCATCACTGCACTCCAGCCTAGTGAACAGAGTGAAGACTCTTGTCAAAA 11983
Qy 1022 AAAAAAAAAA 1033
Db 11984 AAAAAAAAAA 11995

RESULT 15
US-09-949-016-45905/c
; Sequence 45905, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 45905
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-45905

Query Match 18.9%; Score 196.4; DB 4; Length 601;
Best Local Similarity 79.5%; Pred. No. 1.2e-38;
Matches 248; Conservative 1; Mismatches 52; Indels 11; Gaps 1;

Qy 733 TTTTATGTCAGAGGCTGGGCGCATAGCTGATGCTGNGGTTCCAGTGTCTTGCGGGGCC 792
Db 579 TTATATTTTATGAGCCGGGCGATGCTGCTCAGCGCTTATCTCCAGACCTTCGGAGGCT 520
Qy 793 GAGGAGGTGATCACTTAAAGTTCAGAGTTTGAACAGCCTTCCCAACATGTTGAAC 852
Db 519 GAGGTAGCGGATCACTGAGGTTCAGAGTTTCAGACAGCTTGGCAACATGTTGAAT 460
Qy 853 GTTGTCTCTCTAAATAATTAACAGAGCGGTGGTGGACACATCTGTATTCGA 912
Db 459 CCGTCCCTACTTAAATAATTAACAGAGTATGTCGATGTCCTGTATTCGA 400
Qy 913 GCTACTGAGAGGCTTACAGAGAAATTCCTTGAACCTGGAGGAGAGGTTGCACTGA 972
Db 399 GCTACTGAGAGGCTTACAGAGAAATTCCTTGAACCTGGAGGAGAGGTTGCACTGA 340
Qy 973 GC-----CATTCGACTCCAGCCTGGCAACACAGTGAAGCTCTTGTCAAAA 1021
Db 339 GCGGAGATCGCATCACTGCACTCCAGCCTAGTGAACAGAGTGAAGACTCTTGTCAAAA 280
Qy 1022 AAAAAAAAAA 1033
Db 279 AAAAAAAAAA 268
```

Search completed: March 6, 2005, 06:42:04
Job time : 224 secs

CC human secreted proteins given in AAY87064 to AAY87223. Human secreted
CC protein can have activities based on the tissues and cells the genes are
CC expressed in. Examples of activities include: cytostatic; and
CC immunosuppressive; antiinflammatory; nootropic; neuroprotective; and
CC antiallergic. The polynucleotides and their corresponding secreted
CC polypeptides are useful for preventing, treating or ameliorating medical
CC conditions, e.g. by protein or gene therapy. Also pathological conditions
CC can be diagnosed by determining the amount of the new polypeptides in a
CC sample or by determining the presence of mutations in the new
CC polynucleotides. Human secreted protein s and their polynucleotides can
CC be used for developing products for the diagnosis or treatment of cancer,
CC tumours, neurodegenerative disorders, developmental abnormalities and
CC foetal deficiencies, blood disorders, diseases of the immune system,
CC autoimmune diseases, hepatic and renal disease, inflammation, allergies,
CC Alzheimer's disease, behavioural disorders, schizophrenia, osteoporosis,
CC arthritis, infections, AIDS, spinal cord injuries, transplant rejection,
CC diabetes, asthma, sepsis, acne, psoriasis, cardiovascular disorders,
CC reproductive disorders, gastrointestinal disorders, respiratory disorders
CC and metabolic disorders. The proteins or polynucleotides can also be used
CC as food additives or preservatives. The proteins are also useful for
CC identifying their binding partners. AAZ86008 to AAZ98016 and AAY87063 are
CC sequence used in the exemplification of the present invention

CC
XX Sequence 49 AA;

Query Match 100.0%; Score 260; DB 3; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.9e-30; Indels 0; Gaps 0;
Matches 49; Conservative 0; Mismatches 0;

Qy 1 MNLIGMIFSMCGMLKIKWCAMVAVYCSFISFANSRSSEDTKOMSSFM 49
Db 1 MNLIGMIFSMCGMLKIKWCAMVAVYCSFISFANSRSSEDTKOMSSFM 49

RESULT 2
AAE06118
ID AAE06118 standard; protein; 49 AA.

XX
AC AAE06118;

XX
DT 24-SEP-2001 (first entry)

XX
DE Human gene 14 encoded secreted protein HAU183, SEQ ID NO:180.

XX Human, secreted protein; proliferative disorder; cancer; tumour; asthma;
XX foetal abnormality; developmental abnormality; haematopoietic disorder;
XX immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
XX Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
XX psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
XX inflammation; neurological disorder; Alzheimer's disease; food additive;
XX angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;
XX pregnancy-related disorder; endocrine disorder; infection; wound healing;
XX cell culture; chemotaxis; vulnery; binding partner identification;
XX gene therapy; chromosome 19.

XX
OS Homo sapiens.

XX
FH Key Location/Qualifiers

FT Peptide 1..37

FT Protein /label= signal_peptide

FT /label= Mature_human_secreted_protein

XX
PN MO200151504-A1.

XX
PD 19-JUL-2001.

XX
PF 12-JAN-2001; 2001WO-US00911.

XX
PR 13-JAN-2000; 2000US-00482273.

XX
PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Ruben SM, Komatsoulis GA, Duan DR, Rosen CA, Moore PA, Shi Y;
PI Lefleur DM, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR;
PI Endress GA, Muscenski M, Edner R;
XX
XX WPI; 2001-425865/45.

DR N-PSDB; AAD11707.
XX
XX Isolated nucleic acid molecule encoding a human secreted protein is used
PT in preventing, treating or ameliorating a medical condition.

XX
PS Claim 11; Page 796; 864pp; English.

XX
XX AAD11630-AAD11721 represent CDNA's corresponding to 71 human secreted
CC protein genes, and AAE06041-AAE06132 represent the proteins they encode.
CC AAE06133-AAE06205 represent human secreted protein fragments. The
CC secreted proteins and their genes are useful for preventing, treating or
CC ameliorating medical conditions, e.g., by protein or gene therapy.
CC Pathological conditions can be diagnosed by determining the amount of the
CC new protein in a sample or by determining the presence of mutations in
CC the new genes. Specific uses are described for each of the 71 genes,
CC based on the tissues in which they are most highly expressed, and include
CC developing products for the diagnosis or treatment of proliferative
CC disorders, cancer, tumours, foetal and developmental abnormalities,
CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,
CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
CC angiogenic disorders, kidney disorders, gastrointestinal disorders,
CC pregnancy-related disorders, endocrine disorders, and infections. The
CC proteins can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues,
CC to regenerate tissues, to identify their cognate ligands or binding
CC partners, and in chemotaxis, and can be used as a food additive or
CC preservative to modify storage properties. Antibodies specific for a
CC protein of the invention can be used in alleviating symptoms associated
CC with the disorders mentioned above, and in diagnostic immunoassays (e.g.,
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
CC present sequence represents a human secreted protein of the invention

XX
SQ Sequence 49 AA;

Query Match 100.0%; Score 260; DB 4; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.9e-30; Indels 0; Gaps 0;
Matches 49; Conservative 0; Mismatches 0;

Qy 1 MNLIGMIFSMCGMLKIKWCAMVAVYCSFISFANSRSSEDTKOMSSFM 49
Db 1 MNLIGMIFSMCGMLKIKWCAMVAVYCSFISFANSRSSEDTKOMSSFM 49

RESULT 3
ABG33940
ID ABG33940 standard; protein; 49 AA.

XX
AC ABG33940;

XX
DT 15-JUL-2002 (first entry)

XX
DE Human secreted protein encoded by gene 14 #2.

XX Human, secreted protein; gene therapy; immunosuppressive; antiarthritic;
XX antirheumatic; antiproliferative; cytostatic; cardiac; vasotropic;
XX cerebroprotective; nootropic; neuroprotective; antiinfective; virucide;
XX fungicide; ophthalmological; autoimmune disease; neoplasm;
XX rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;
XX cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
XX angiogenesis; nervous system disorder; Alzheimer's disease; infection;
XX ocular disorder; corneal infection; wound healing; skin aging;
XX epithelial cell proliferation; food additive.

XX
OS Homo sapiens.

XX WO200226931-A2.
PN
XX
XX 04-APR-2002.
PD
XX
XX 24-SEP-2001; 2001WO-US029871.
PP
XX 25-SEP-2000; 2000US-0234925P.
PR
XX 12-JAN-2001; 2001WO-US000911.
PR
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Ruben SM, Komatsoulis G, Duan DR, Rosen CA, Moore PA, Shi Y,
PI Lafleur DM, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR,
PI Endress GA, Muceneki M, Ebner R;
XX WPI; 2002-362489/39.
DR N-PSDB; ABK69803.
XX
XX Novel 71 isolated secreted polypeptides and polynucleotides encoding the
PT polypeptides, useful for treating Huntington's disease, sepsis,
PT meningitis, thrombocytopenia, hemolytic anemia, rheumatoid arthritis,
PT asthma.
XX
XX Claim 11; Page 1273; 1478pp; English.
PS
XX The invention relates to an isolated nucleic acid molecule (or its
CC fragment, homologue complement or allelic variant) encoding a human
CC secreted protein (and its fragment, domain, epitope, variant, secreted
CC form and species variant). Also included are a recombinant vector
CC comprising the nucleic acid, a recombinant host cell comprising the
CC vector, an antibody against the secreted protein, a recombinant host cell
CC that expresses the secreted protein and a method of identifying a binding
CC partner of the secreted protein. The nucleic acid and protein are used to
CC prevent, diagnose, treat or ameliorate a medical condition in e.g.
CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep for
CC example autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia,
CC angiodysplasia, nervous system disorders e.g. Alzheimer's disease,
CC infections caused by bacteria, viruses and fungi and ocular disorders
CC e.g. corneal infection. Many other diseases and disorders are listed in
CC the specification. The polypeptides can also be used to aid wound healing
CC an epithelial cell proliferation, to prevent skin aging due to sunburn, of
CC to maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. The present sequence
CC represents a novel human secreted protein of the invention
XX
SQ Sequence 49 AA;
Query Match 100.0%; Score 260; DB 5; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.9e-30;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMCGMLTKIKWCANVAVYCSFISFANSRSSEDTKOMSSFM 49
Db 1 MNLIGMIFSMCGMLTKIKWCANVAVYCSFISFANSRSSEDTKOMSSFM 49

RESULT 4
ABR47655
ID ABR47655 standard; protein; 49 AA.
XX
XX ABR47655;
AC
XX
XX 12-JUN-2003 (first entry)
DT
XX
XX Human secreted protein, SEQ ID 546.
DE
XX
XX Cardiac; antiarrhythmic; antiarteriosclerotic; vasotropic; cyostatic;
KW vulnerable; antiinflammatory; nootropic; neuroprotective;

KW antiparkinsonian; gene therapy; human; cardiovascular disorder.
XX
XX Homo sapiens.
OS
XX
XX WO200295010-A2.
FN
XX
XX 28-NOV-2002.
PD
XX
XX 19-MAR-2002; 2002WO-US009785.
PP
XX
XX 21-MAR-2001; 2001US-0277340P.
PR
XX 19-JUL-2001; 2001US-0306171P.
PR
XX 13-NOV-2001; 2001US-0331287P.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Rosen CA, Ruben SM;
PI
XX WPI; 2003-129429/12.
DR
XX
XX Novel human secreted proteins, useful for detecting, preventing,
PT diagnosing, prognosticating, treating and/or ameliorating cardiovascular
PT disorders such as arrhythmia.
XX
XX Claim 13; SEQ ID NO 546; 1881pp; English.
PS
XX
XX The present invention relates to novel human secreted proteins (ABR47653-
CC ABR48145) and their coding sequences (ACC50344-ACC50586). The proteins
CC and their coding sequences are useful for the preparation of a diagnostic
CC or pharmaceutical composition for diagnosing or treating a cardiovascular
CC disorder (e.g., arrhythmia, tachycardia, cardiac arrest, coronary
CC arteriosclerosis and myocardial ischemia), neural disorders, immune
CC system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, renal disorders,
CC proliferative disorders and/or cancerous diseases and conditions, for
CC wound healing and epithelial cell proliferation, to treat inflammation or
CC infection, for treating thrombosis and arteriosclerosis, for treating or
CC preventing neural damage which occurs in neuronal disorders or
CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's
CC disease, to enhance bone and periodontal regeneration and aid in tissue
CC transplants or bone grafts, to prevent skin aging or hair loss, to
CC stimulate growth and differentiation of haematopoietic cells and bone
CC marrow cells when used in combination with other cytokines, to maintain
CC organs before transplantation or for supporting cell culture of primary
CC tissues, to increase or decrease differentiation or proliferation of
CC embryonic stem cells, or to modulate mammalian characteristics or
CC metabolism. Note: The sequence data for this patent was published in
CC electronic format and is available from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 49 AA;
Query Match 100.0%; Score 260; DB 6; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.9e-30;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMCGMLTKIKWCANVAVYCSFISFANSRSSEDTKOMSSFM 49
Db 1 MNLIGMIFSMCGMLTKIKWCANVAVYCSFISFANSRSSEDTKOMSSFM 49

RESULT 5
ABR00022
ID ABR00022 standard; protein; 49 AA.
XX
XX ABR00022;
AC
XX
XX 03-APR-2003 (first entry)
DT
XX
XX Human gene 12 encoded secreted protein HAVUA183, SEQ ID NO:311.
DE
XX
XX Human; secreted protein; digestive disorder; gastrointestinal disorder;
KW mouth; oesophagus; stomach; small intestine; large intestine; liver;

KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
KW immune disorder; inflammation; infection; wound healing; drug screening;
KW chromosome identification; chromosome mapping; cytostatic; gene therapy;
KW antiinflammatory; immunosuppressive; vulnery; chromosome 19.
OS Homo sapiens.
XX
XX WO200276488-A1.
PN
XX
PD 03-OCT-2002.
XX
PF 19-MAR-2002; 2002WO-US008276.
XX
XX
XX 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
PI
XX
XX WPI; 2003-029900/02.
DR N-PSDB; ABZ71201.
XX
XX
PT New human secreted proteins and nucleic acids, useful for detecting,
PT preventing, diagnosing, prognosticating, treating and/or ameliorating
PT e.g. gastrointestinal diseases and disorders, or cancers.
XX
XX
PS Claim 13; Page 942; 1216pp; English.
XX
XX
CC ABZ71190-ABZ71478 represent cDNAs corresponding to 178 human secreted
CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
CC ABZ71479-ABZ71540 represent human secreted protein genomic fragments. The
CC invention also encompasses antibodies specific for the secreted proteins,
CC the use of the secreted proteins in drug screening and recombinant
CC vectors and host cells comprising a nucleic acid of the invention. The
CC secreted proteins, nucleic acids encoding them, antibodies or antibody
CC fragments specific for the secreted proteins, and modulators of protein
CC activity are useful for diagnosing, treating, ameliorating or preventing
CC digestive disorders. Such conditions include disorders of the mouth,
CC oesophagus, stomach, small intestine, large intestine, liver, biliary
CC tract and pancreas, and include cancers of these organs and tissues. The
CC secreted proteins and their nucleic acids may also be used in the
CC treatment of immune disorders, inflammation, infection,
CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
CC of the invention may be used for chromosome identification, chromosome
CC mapping, in gene therapy, for identifying individuals from minute
CC biological samples, as hybridisation probes, and as molecular weight
CC markers. The present sequence represents a human secreted protein of the
CC invention
XX
XX
SQ Sequence 49 AA;
Query Match 100.0%; Score 260; DB 6; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.9e-30;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNLLGMIFSMGMLKLKMKCAWVAAYCSFISFANSRSSEDTKQMMSSFM 49
Db 1 MNLLGMIFSMGMLKLKMKCAWVAAYCSFISFANSRSSEDTKQMMSSFM 49
RESULT 6
ADB91468
ID ADB91468 standard; protein; 49 AA.
XX
AC ADB91468;
XX
XX
DT 04-DEC-2003 (first entry)
XX
DE Human secreted protein #SEQ ID 414.
XX
KW Secreted protein; gene therapy; antidiabetic; diabetes; human.

XX
OS Homo sapiens.
XX
XX WO2003004622-A2.
PN
XX
XX 16-JAN-2003.
XX
XX
XX 19-MAR-2002; 2002WO-US008124.
PF
XX
XX 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Rosen CA, Ruben SM;
PI
XX
XX WPI; 2003-229407/22.
DR
XX
XX
PT Nucleic acid encoding a human secreted protein is useful in diagnosing or
PT treating diabetes or conditions related to diabetes.
XX
XX
PS Claim 3; SEQ ID NO 414; 1537pp; English.
XX
XX
CC The invention relates to isolated nucleic acid molecules ADB91065-
CC ADB91448 and ADB91835-ADB91911 encoding human secreted proteins ADB91449-
CC ADB91834. Also disclosed is a recombinant vector comprising a
CC polynucleotide of the invention, and a recombinant host cell comprising
CC the recombinant vector. The polypeptide of the invention is useful in
CC identifying a binding partner by contacting the polypeptide with a
CC binding partner, and determining whether the binding partner increases or
CC decreases activity of the polypeptide. The polypeptide, polynucleotide,
CC antibody or its fragment, agonist or antagonist are useful for preparing
CC a pharmaceutical composition for diagnosing or treating diabetes or
CC conditions related to diabetes. The present sequence is that of the human
CC immunoglobulin Fc portion used to generate fusion proteins, increasing
CC the stability of the fused protein as compared to the secreted protein
CC only. Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
XX
XX
SQ Sequence 49 AA;
Query Match 100.0%; Score 260; DB 7; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.9e-30;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNLLGMIFSMGMLKLKMKCAWVAAYCSFISFANSRSSEDTKQMMSSFM 49
Db 1 MNLLGMIFSMGMLKLKMKCAWVAAYCSFISFANSRSSEDTKQMMSSFM 49
RESULT 7
ADCT74024
ID ADCT74024 standard; protein; 49 AA.
XX
XX
AC ADCT74024;
XX
XX
DT 01-JAN-2004 (first entry)
XX
XX
DE Human secreted protein - SEQ ID 657.
XX
XX
KW antianemic; antirheumatic; antiarthritic; antiinflammatory; antithyroid;
KW antidiabetic; immunosuppressive; dermatological; nephrotoxic;
KW antiparkinsonian; neuroprotective; nootropic; antibacterial; virostatic;
KW fungicide; antiparasitic; antiarteriosclerotic; vulnery; cytostatic;
KW haemopoietic; haematologic; anaemia; autoimmune disorder;
KW rheumatoid arthritis; inflammation; Grave's disease; diabetes;
KW systemic lupus erythematosus; glomerulonephritis; neurodegenerative;
KW Parkinson's; Alzheimer's; wound; hyperproliferative; atherosclerosis;
KW cancer; bacterial; viral; fungal; parasitic infection; gene therapy;
XX
XX human.

OS Homo sapiens.
XX
PN WO2003038663-A2.
XX
PD 08-MAY-2003.
XX
PF 19-MAR-2002; 2002WO-US008277.
XX
PR 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI: 2003-430516/40.
DR N-PSDB; ADC73409.
XX
XX New human secreted polypeptide for diagnosing, preventing or treating
PT hematopoietic or hematologic disorders (e.g. anemia), autoimmune
PT disorders (e.g. diabetes) or hyperproliferative disorders (e.g. cancer or
PT atherosclerosis).
XX
PS Claim 16; SEQ ID NO 657; 2272bp; English.
XX
XX The invention relates to a novel human secreted polypeptide comprising a
CC defined sequence given in the specification. The polypeptide, nucleic
CC acid molecule, antibody, agonist or antagonist of the invention may be
CC useful for preparing a composition for diagnosing or treating a
CC hematopoietic or haematologic disorder such as anaemia, autoimmune
CC disorders such as rheumatoid arthritis, inflammation, Grave's disease,
CC diabetes, systemic lupus erythematosus or glomerulonephritis,
CC neurodegenerative disorders including Parkinson's disease and Alzheimer's
CC disease, wounds and hyperproliferative disorders including
CC atherosclerosis or cancer, as well as bacterial, viral, fungal or
CC parasitic infections. The polypeptide may also be used during gene
CC therapy procedures and for identifying a binding partner by contacting
CC the polypeptide with a binding partner and determining whether the
CC binding partner increases or decreases the activity of the polypeptide.
CC The current sequence is that of the human secreted protein of the
CC invention.
XX
SQ Sequence 49 AA;
Query Match 100.0%; Score 260; DB 7; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.3e-30;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNLIGMIFSMGMLKLKMCAMVAVYCSFISFANSRSSEDTKOMSSFM 49
DB 1 MNLIGMIFSMGMLKLKMCAMVAVYCSFISFANSRSSEDTKOMSSFM 49
RESULT 8
AAW67873 ID AAW67873 standard; protein; 50 AA.
XX
AC AAW67873;
XX
DT 25-MAR-1999 (first entry)
XX
DE Human secreted protein.
XX
KM Human; secreted protein; fusion protein; gene therapy; protein therapy;
KM diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KM developmental abnormality; foetal deficiency; blood; allergy; renal;
KM immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KM inflammation; ischaemic shock; Alzheimer's disease; resveratrol; AIDS;
KM cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KM osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KM endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX

OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 50
FT /Label= unknown
XX
XX WO9842738-A1.
XX
XX 01-OCT-1998.
XX
XX 19-MAR-1998; 98WO-US005311.
XX
XX 21-MAR-1997; 97US-0041276P.
XX 21-MAR-1997; 97US-0041277P.
XX 21-MAR-1997; 97US-0041281P.
XX 21-MAR-1997; 97US-0042344P.
XX 30-MAY-1997; 97US-0048069P.
XX 30-MAY-1997; 97US-0048094P.
XX 30-MAY-1997; 97US-0048095P.
XX 30-MAY-1997; 97US-0048096P.
XX 30-MAY-1997; 97US-0048099P.
XX 30-MAY-1997; 97US-0048131P.
XX 30-MAY-1997; 97US-0048135P.
XX 30-MAY-1997; 97US-0048154P.
XX 30-MAY-1997; 97US-0048160P.
XX 30-MAY-1997; 97US-0048166P.
XX 30-MAY-1997; 97US-0048187P.
XX 30-MAY-1997; 97US-0048188P.
XX 30-MAY-1997; 97US-0048351P.
XX 30-MAY-1997; 97US-0048352P.
XX 30-MAY-1997; 97US-0048353P.
XX 30-MAY-1997; 97US-0048355P.
XX 30-MAY-1997; 97US-0050937P.
XX 05-AUG-1997; 97US-0054804P.
XX 19-AUG-1997; 97US-0056370P.
XX 02-OCT-1997; 97US-0060862P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Young P, Greene JM, Ferrie AM, Ruben SM, Rosen CA, Duan R, Hu J;
XX Florence KA, Olsen HS, Ebner R, Brewer LA, Moore PA, Shi Y;
XX Lafleur DW, Ni J;
XX
XX WPI: 1999-070066/06.
XX DR N-PSDB; AAX00677.
XX
XX New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders.
XX
XX Claim 11; Page 308; 385pp; English.
XX
XX This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. AAX00602) for increasing the stability of the fused protein
CC as compared to the human protein only. The invention relates to 87 novel
CC genes and their fragments (nucleic acid sequences: AAX00611-X00724; amino
CC acid sequences AAW67807-W68004) which are useful for preventing, treating
CC or ameliorating medical conditions e.g. by protein or gene therapy. Also,
CC pathological conditions can be diagnosed by determining the amount of the
CC new polypeptides in a sample or by determining the presence of mutations
CC in the new polynucleotides. Specific uses are described for each of the
CC 87 polynucleotides, based on which tissues they are most highly expressed
CC in (see AAX00611 for described uses)
XX
SQ Sequence 50 AA;
Query Match 100.0%; Score 260; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 3e-30;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNLIGMIFSMGMLKLKMCAMVAVYCSFISFANSRSSEDTKOMSSFM 49

Db 1 MNLIGMIFSMCGMLKLMKCAWVAVYCSFISFANSRSSSEDTKOMSSFM 49

RESULT 9
AA76058

ID AAY76058 standard; protein; 70 AA.

AC AAY76058;

DT 27-MAR-2000 (first entry)

DE Human skin cell protein, SEQ ID NO:313.

XX Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
KW secreted; transmembrane; inflammation; cancer; neurological disease;
KW angiogenesis; tumor vasculatization; growth disorder;
KW developmental disorder; skin wound; hair follicle disorder;
KW anti-inflammatory; cytostatic; neuroprotective; vulnery.

OS Homo sapiens.

PN WO955865-A1.

PD 04-NOV-1999.

PF 29-APR-1999; 99WO-NZ000051.

PR 29-APR-1998; 98US-00069726.

PR 09-NOV-1998; 98US-00188930.

PA (GENE-) GENESIS RES & DEV CORP LTD.

PI Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;

DR WPI; 2000-072177/06.

PT Novel polynucleotides useful for the treatment of various conditions
including wounds and cancer.

PS Claim 4; Page 185; 235pp; English.

XX The invention relates to novel nucleic acid sequences derived from rat
XX dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
XX and mouse embryonic skin, keratinocyte stem cells and transit amplifying
XX cells. Polypeptides of the invention may be used to treat inflammation,
XX cancer and neurological diseases. The proteins may be used to stimulate
XX the growth and motility of keratinocytes, to inhibit the growth of cancer
XX cells, to modulate angiogenesis and tumor vasculatization, to modulate
XX skin inflammation, to modulate epithelial cell growth and to inhibit
XX binding of HIV-1 to leukocytes. The invention may also be used to treat
XX growth and developmental defects, skin wounds and hair follicle
XX disorders. Sequences AAY75942-Y76123 represent polypeptides encoded by
XX cDNA sequences derived from several mouse, rat or human skin cell types.
XX Sequences AAY75942-Y75947, AAY76020-Y76021, AAY76094-Y76104 and AAY76119
XX are proteins with an N-terminal signal sequence, indicating that they are
XX secreted. Sequences AAY75986-Y75989, AAY76061-Y76071, AAY76106-Y76109 and
XX AAY76121-Y76122 are proteins with one or more putative transmembrane
XX domains

SO Sequence 70 AA;

Query Match 100.0%; Score 260; DB 3; Length 70;

Best Local Similarity 100.0%; Pred. No. 4.3e-30;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMCGMLKLMKCAWVAVYCSFISFANSRSSSEDTKOMSSFM 49

Db 1 MNLIGMIFSMCGMLKLMKCAWVAVYCSFISFANSRSSSEDTKOMSSFM 49

RESULT 10
AAB55997

ID AAB55997 standard; protein; 70 AA.

AC AAB55997;

DT 08-MAR-2001 (first entry)

DE Skin cell protein, SEQ ID NO: 313.

XX Human; skin cell; cytostatic; antiinflammatory; anti-HIV; nootropic;
KW neuroprotective; vulnery; immunomodulatory; vaccine;
KW keratinocyte growth stimulation; cancer; angiogenesis inhibition;
KW inflammation; neurological disease.

OS Homo sapiens.

PN WO200069884-A2.

PD 23-NOV-2000.

PF 15-MAY-2000; 2000WO-NZ000075.

PR 14-MAY-1999; 99US-00312283.

PA (GENE-) GENESIS RES & DEV CORP LTD.

PI Watson JD, Strachan L, Onrust R, Sleeman M, Kumble KD;

DR WPI; 2001-007495/01.

DR N-PSDB; AAC99696.

XX New isolated polynucleotide used in the identification of genetic
XX disorders and encoding polypeptides used for treating inflammatory
XX disease, cancer and neurological diseases.

PS Claim 4; Page 251; 352pp; English.

XX The present sequence is a polypeptide which is expressed in mammalian
XX skin cells. The polypeptide is useful for stimulating keratinocyte growth
XX and motility, inhibiting the growth of cancer cells, modulating
XX angiogenesis, inhibiting angiogenesis and vasculatization of tumours,
XX modulating skin inflammation, stimulating the growth of epithelial cells,
XX inhibiting the binding of human immunodeficiency virus (HIV)-1 to
XX leukocytes, and treating inflammatory disease, cancer and neurological
XX diseases. The polynucleotide can be used as a marker, in the
XX identification of genetic disorders, and for the design of
XX oligonucleotides for examining expression patterns

SO Sequence 70 AA;

Query Match 100.0%; Score 260; DB 4; Length 70;

Best Local Similarity 100.0%; Pred. No. 4.3e-30;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMCGMLKLMKCAWVAVYCSFISFANSRSSSEDTKOMSSFM 49

Db 1 MNLIGMIFSMCGMLKLMKCAWVAVYCSFISFANSRSSSEDTKOMSSFM 49

RESULT 11

ID AAE06054 standard; protein; 70 AA.

AC AAE06054;

DT 24-SEP-2001 (first entry)

DE Human gene 14 encoded secreted protein HHEBQ41, SEQ ID NO:116.

XX Human; secreted protein; proliferative disorder; cancer; tumour; asthma;
KW foetal abnormality; developmental abnormality; hematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;

XX XX Homo sapiens.

XX XX Key Location/Qualifiers

XX XX Peptide 1..35

XX XX Protein 36..70

XX XX /label= Mature_human_secreted_protein

XX XX WO200151504-A1.

XX XX 19-JUL-2001.

XX XX 12-JAN-2001; 2001WO-US000911.

XX XX 13-JAN-2000; 2000US-00482273.

XX XX (HUMA-) HUMAN GENOME SCI INC.

XX XX Ruben SM, Komatsu J GA, Duan DR, Rosen CA, Moore PA, Shi Y,

P1 Lafleur DW, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR,

P1 Endress GA, Musceni M, Ebner R;

XX XX MPI; 2001-425865/45.

DR N-PSDB: AAD11643.

XX XX

XX XX Isolated nucleic acid molecule encoding a human secreted protein is used

PT in preventing, treating or ameliorating a medical condition.

PS Claim 11; Page 747; 864pp; English.

XX XX

XX XX AAD11630-AAD11721 represent cDNAs corresponding to 71 human secreted

CC protein genes, and AAE06041-AAE06132 represent the proteins they encode.

CC AAE06133-AAE06205 represent human secreted protein fragments. The

CC secreted proteins and their genes are useful for preventing, treating or

CC ameliorating medical conditions, e.g., by protein or gene therapy.

CC Pathological conditions can be diagnosed by determining the amount of the

CC new protein in a sample or by determining the presence of mutations in

CC the new genes. Specific uses are described for each of the 71 genes,

CC based on the tissues in which they are most highly expressed, and include

CC developing products for the diagnosis or treatment of proliferative

CC disorders, cancer, tumors, foetal and developmental abnormalities,

CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune

CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,

CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),

CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,

CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,

CC angiogenic disorders, kidney disorders, gastrointestinal disorders,

CC pregnancy-related disorders, endocrine disorders, and infections. The

CC proteins can also be used to aid wound healing and epithelial cell

CC proliferation, to prevent skin aging due to sunburn, to maintain organs

CC before transplantation, for supporting cell culture of primary tissues,

CC to regenerate tissues, to identify their cognate ligands or binding

CC partners, and in chemotaxis, and can be used as a food additive or

CC preservative to modify storage properties. Antibodies specific for a

CC protein of the invention can be used in alleviating symptoms associated

CC with the disorders mentioned above, and in diagnostic immunoassays (e.g.,

CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The

CC present sequence represents a human secreted protein of the invention

XX XX

XX XX Sequence 70 AA;

XX XX

XX XX Query Match 100.0%; Score 260; DB 4; Length 70;

XX XX Best Local Similarity 100.0%; Pred. No. 4,3e-30;

XX XX Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0

XX XX

XX XX 1 NMLLGIMFISMGCLMLKMCAMVAVYCSFISFANSRSSEDTQKMMSSFM 49

DB	1	MMLGITSMGGLMKMKCANVAVYCSFISFANRSSEDTQMSSFM	49
RESULT 12			
ABG33876			
ID	ABG33876	standard; protein; 70 AA.	
XX			
AC	ABG33876;		
XX			
DT	15-JUL-2002	(first entry)	
XX			
DE	Human secreted protein encoded by gene 14 #1.		
XX			
KM	Human; secreted protein; gene therapy; immunosuppressive; antiarthritic;		
KM	antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;		
KM	carcinoprotective; nootropic; neuroprotective; antibacterial; vitruicide;		
KM	fungicide; ophthalmological; autoimmune disease; neoplasm;		
KM	rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;		
KM	cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;		
KM	angiogenesis; nervous system disorder; Alzheimer's disease; infection;		
KM	ocular disorder; corneal infection; wound healing; skin aging;		
XX	epithelial cell proliferation; food additive.		
OS	Homo sapiens.		
XX			
PN	WO20026931-A2.		
XX			
PD	04-APR-2002.		
XX			
PF	24-SEP-2001; 2001WO-US029871.		
XX			
PR	25-SEP-2000; 2000US-0234925P.		
PR	12-JAN-2001; 2001WO-US000911.		
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
XX			
PI	Ruben SM, Komatsoulis G, Duan DR, Rosen CA, Moore PA, Shi Y,		
PI	Lafleur DW, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR;		
PI	Endresse GA, Mucenski M, Ebner R;		
XX			
DR	WPI; 2002-362489/39.		
XX			
DR	N-PSDB; ABR69739.		
XX			
PT	Novel 71 isolated secreted polypeptides and polynucleotides encoding the		
PT	polypeptides, useful for treating Huntington's disease, sepsis,		
PT	meningitis, thrombocytopenia, hemolytic anemia, rheumatoid arthritis,		
PT	asthma.		
XX			
PS	Claim 11; Page 1228; 1478pp; English.		
XX			
CC	The invention relates to an isolated nucleic acid molecule (or its		
CC	fragment, homologue complement or allelic variant) encoding a human		
CC	secreted protein (and its fragment, domain, epitope, variant, secreted		
CC	form and species variant). Also included are a recombinant vector		
CC	comprising the nucleic acid, a recombinant host cell comprising the		
CC	vector, an antibody against the secreted protein, a recombinant host cell		
CC	that expresses the secreted protein and a method of identifying a binding		
CC	partner of the secreted protein. The nucleic acid and protein are used to		
CC	prevent, diagnose, treat or ameliorate a medical condition in e.g.		
CC	humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep for		
CC	example autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative		
CC	disorders e.g. neoplasms of the breast or liver, cardiovascular disorders		
CC	e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,		
CC	angiogenesis, nervous system disorders e.g. Alzheimer's disease,		
CC	infections caused by bacteria, viruses and fungi and ocular disorders		
CC	e.g. corneal infection. Many other diseases and disorders are listed in		
CC	the specification. The polypeptides can also be used to aid wound healing		
CC	an epithelial cell proliferation, to prevent skin aging due to sunburn,		
CC	to maintain organs before transplantation, for supporting cell culture of		
CC	primary tissues, to regenerate tissues and in chemotaxis. The		
CC	polypeptides can also be used as a food additive or preservative to		
CC	increase or decrease storage capabilities. The present sequence		

CC represents a novel human secreted protein of the invention
 XX
 SQ Sequence 70 AA;

Query Match 100.0%; Score 260; DB 5; Length 70;
 Best local Similarity 100.0%; Pred. No. 4.3e-30;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMCGMLTKIKMCAMWAVYCSFISFANSRSSEDTKOMSSFM 49
 1 MNLIGMIFSMCGMLTKIKMCAMWAVYCSFISFANSRSSEDTKOMSSFM 49
 Db

RESULT 13

AB872197 standard; protein: 70 AA.

AC AB872197;

XX 04-APR-2002 (first entry)

DT Human protein isolated from skin cells SEQ ID NO: 313.

XX Human; rat; mouse; skin cell; skin wound; cancer; growth defect;
 KW developmental defect; inflammatory disease; dermatological; vulnary;
 KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.

XX Homo sapiens.

OS WO200190357-A1.

XX 29-NOV-2001.

PD 24-MAY-2001; 2001WO-NZ000099.

PR 24-MAY-2000; 2000US-0206650P.

PR 25-JUL-2000; 2000US-0221232P.

XX (GENE-) GENESIS RES & DEV CORP LTD.

PI Watson JD, Strachan L, Sleeman M, Onrust R, Morrison JG;

XX Kumble KD;

DR WPI; 2002-122020/16.

XX New polynucleotides and polypeptides encoded by the polynucleotides
 PT isolated from skin cells, useful for treating skin wounds, cancers,
 PT growth and developmental defects, inflammatory diseases, or for
 PT modulating immune responses.

XX Example 2; Page 207; 466pp; English.

XX The present invention provides the protein and coding sequences of cDNAs
 CC isolated from human, murine and rat skin cell libraries. The sequences
 CC can be used in the development of therapeutic agents useful in the
 CC treatment of skin diseases, including skin wounds, cancer, growth
 CC defects, developmental defects and inflammatory diseases. The proteins
 CC have important roles in the induction of hair growth, cell proliferation
 CC and cell-cell interaction, in maintaining tissue integrity, in wound
 CC healing and in modulating immune responses. The present sequence is a
 CC polypeptide of the invention

XX Sequence 70 AA;

Query Match 100.0%; Score 260; DB 5; Length 70;
 Best local Similarity 100.0%; Pred. No. 4.3e-30;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMCGMLTKIKMCAMWAVYCSFISFANSRSSEDTKOMSSFM 49
 1 MNLIGMIFSMCGMLTKIKMCAMWAVYCSFISFANSRSSEDTKOMSSFM 49
 Db

RESULT 14
 ID AAY87077 standard; protein: 71 AA.

AC AAY87077;

DT 09-MAY-2000 (first entry)

XX Human secreted protein sequence SEQ ID NO:116.

XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
 KW antiinflammatory; neurotropic; neuroprotective; antiallergic; cancer;
 KW tumor; neurodegenerative disorder; developmental abnormality; allergy;
 KW foetal deficiency; blood disorder; immune system disorder; arthritis;
 KW autoimmune disease; hepatic disease; renal disease; inflammation;
 KW Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;
 KW infection; AIDS; spinal cord injury; transplant rejection; diabetes;
 KW asthma; sepsis; acne; psoriasis; cardiovascular disorder;
 KW reproductive disorder; gastrointestinal disorder; respiratory disorder;
 KW metabolic disorder; food additive; preservative; chromosome 19.

XX Homo sapiens.

PN WO200004140-A1.

PD 27-JAN-2000.

PF 14-JUL-1999; 99WO-US015849.

PR 15-JUL-1998; 98US-0092921P.

PR 15-JUL-1998; 98US-0092922P.

PR 15-JUL-1998; 98US-0092956P.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Komatsu S, G, Duan RD, Rosen CA, Moore PA, Shi Y;

PI Lafleur DW, Edner R, Olsen HS, Brewer LA, Florence KA, Young PE;

PI Mucenski M, Endress GA, Soppet DK;

DR WPI; 2000-161128/14.

DR N-PDB; AAZ98030.

XX New isolated human genes, useful for diagnosis and treatment of, e.g.

XX cancers, neurological or blood disorders.

XX Claim 11; Page 380; 494pp; English.

XX The polynucleotide sequences given in AAZ98017 to AAZ98108 encode the
 CC human secreted proteins given in AAY87064 to AAY87223. Human secreted
 CC protein can have activities based on the tissues and cells the genes are
 CC expressed in. Examples of activities include: cytostatic;
 CC immunosuppressive; antiinflammatory; neurotropic; neuroprotective; and
 CC antiallergic. The polynucleotides and their corresponding secreted
 CC polypeptides are useful for preventing, treating or ameliorating medical
 CC conditions, e.g. by protein or gene therapy. Also pathological conditions
 CC can be diagnosed by determining the amount of the new polypeptides in a
 CC sample or by determining the presence of mutations in the new
 CC polynucleotides. Human secreted proteins and their polynucleotides can
 CC be used for developing products for the diagnosis or treatment of cancer,
 CC tumours, neurodegenerative disorders, developmental abnormalities and
 CC foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, hepatic and renal diseases, inflammation, allergies,
 CC Alzheimer's disease, behavioural disorders, schizophrenia, osteoporosis,
 CC arthritis, infections, AIDS, spinal cord injuries, transplant rejection,
 CC diabetes, asthma, sepsis, acne, psoriasis, cardiovascular disorders,
 CC reproductive disorders, gastrointestinal disorders, respiratory disorders
 CC and metabolic disorders. The proteins or polynucleotides can also be used
 CC as food additives or preservatives. The proteins are also useful for
 CC identifying their binding partners. AAZ98008 to AAZ98016 and AAY87063 are
 CC sequence used in the exemplification of the present invention

XX Sequence 71 AA;

